

AD-A185 398 POINT PROCESSES(U) R F SERFOZO MAY 87 TR-185
AFOSR-TR-87-1105 F49620-85-C-0144

AFOSR-TR-87-1105 F49620-85-C-0144

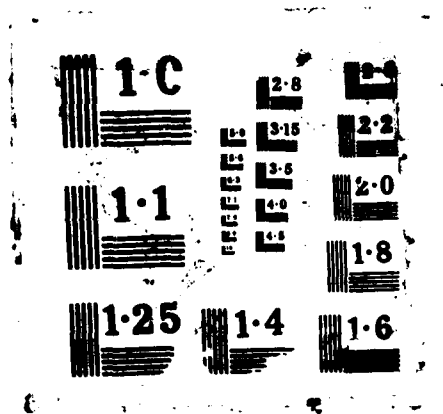
1/1

UNCLASSIFIED

F/G 12/3

NL

117
118



DTIC FILE COPY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A185 398

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY NA		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution Unlimited	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-87-1105	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 185		7a. NAME OF MONITORING ORGANIZATION AFOSR/NM	
6a. NAME OF PERFORMING ORGANIZATION University of North Carolina		7b. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, DC 20332-6448	
6b. OFFICE SYMBOL (If applicable)		7c. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, DC 20332-6448	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620 85 C 0144	
8b. OFFICE SYMBOL (If applicable) NM		10. SOURCE OF FUNDING NOS.	
8c. ADDRESS (City, State and ZIP Code) Building 410, Bolling AFB, DC 20332-6448		PROGRAM ELEMENT NO. 6.1102F	
11. TITLE (Include Security Classification) Point processes		PROJECT NO. 2304	
12. PERSONAL AUTHOR(S) Serfozo, R.F.		TASK NO. A5	
13a. TYPE OF REPORT preprint		14. DATE OF REPORT (Yr., Mo., Day) May 1987	
13b. TIME COVERED FROM 9/86 TO 9/87		15. PAGE COUNT 82	
16. SUPPLEMENTARY NOTATION → This document			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.	
		Keywords and phrases:	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) I shall describe the structure of most of these processes and discuss some of their basic properties. The coverage does not include several important topics requiring lengthy mathematical development (e.g. martingale theory of point processes, general Palm probabilities, and ergodic and spectral analysis of stationary processes). The emphasis will be on presenting tools for modeling stochastic systems rather than on applications of the tools. Although the theory of point processes is intimately connected with the subject of measure and integration (a point process is a random counting measure). I shall focus on results that can be understood without a deep knowledge of measure theory. On the other hand, the presentation will be rigorous and at the level			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Peggy Ravitch		22b. TELEPHONE NUMBER (Include Area Code) 919-962-2307	
22c. OFFICE SYMBOL AFOSR/NM		22d. OFFICE SYMBOL AFOSR/NM	

19 ABSTRACT of the applied probability literature that one would encounter in studying point processes.

To see what topics lie ahead, consult the table of contents.

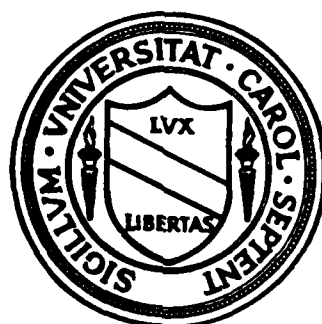
Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



CENTER FOR STOCHASTIC PROCESSES

AFOSR-TR- 87-1105

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina



POINT PROCESSES

by

Richard F. Serfozo

Technical Report No. 185

May 1987

87 5 1 293

POINT PROCESSES

by

Richard F. Serfozo
Georgia Institute of Technology

*These notes constitute a first draft of a chapter for the Operations Research Handbook on Stochastic Processes. April 1978.

**This work was sponsored in part by the Air Force Office of Scientific Research under contracts 84-0367 and F49620 85 C .0144

Contents

	Page
Section 1. Introduction	1
1.1 Literature on Point Processes	1
1.2 Definition of a Point Process	4
1.3 Distributions and Laplace Transforms of Point Processes	8
1.4 Basic Examples: Poisson, Renewal and Stationary Processes	10
1.5 Marked and Compound Point Processes	14
Section 2. Poisson Processes and Some Relatives	18
2.1 Characterization of Poisson Processes	18
2.2 Sample Processes	20
2.3 Sums, Partitions, Thinnings and Translations of Poisson Processes	22
2.4 Cox Processes	26
2.5 Poisson Cluster Processes	29
Section 3. Convergence in Distribution of Point Processes	32
3.1 Basics of Convergence in Distribution	32
3.2 Convergence to Poisson Processes	34
3.3 Convergence to Infinitely Divisible Point Processes	39
3.4 Poisson Approximations	42
Section 4. Renewal Theory	47
4.1 Distributions of Renewal Processes	47
4.2 Key Renewal Theorem	50
4.3 Applications of the Key Renewal Theorem	52
4.4 Laws of Large Numbers	55
Section 5. Stationary Processes	59
5.1 Definitions and Examples	59
5.2 Infinitesimal Properties	61
5.3 Palm Probabilities	64
Section 6. Point Processes Characterized by Martingales	68
6.1 Compensators of Point Processes	69
6.2 Poisson Convergence and Approximations	73
6.3 Customer Flows in a Jackson Network That are Approximately Poisson	74

1. INTRODUCTION

1.1 Literature on Point Processes

Point processes are models of random numbers of events in time intervals or numbers of points in regions. Here are some typical families of examples.

Times at which an event occurs: Times of births, police emergencies, failures of a machine, insurance claims, or earthquakes.

Random flows or streams of items: Times at which items enter or leave a certain place such as telephone calls arriving to a switching center, data packets entering a computer, parts leaving a manufacturing work station, and cash flows in a company.

Random locations of points in an Euclidean space: Galaxies in space, errors in a computer code, animals in a forest, aircraft over a city.

Times of special events in a stochastic process: The instants when a Gaussian process crosses a certain level or when a pure-jump Markov process makes a certain type of transition.

Random locations of elements in an abstract set: One can talk of a point process in which the points are functions in a space of functions, lines in a set of lines on the plane, graphs in a set of graphs, etc.

Although the theory of point processes has been developed only recently, its origins go back several centuries. Here are its major roots.

Poisson Phenomena. Poisson (1837) showed that the Poisson distribution is the limit of a binomial distribution of rare events. This led to numerous applications of the Poisson distribution in the nineteenth century and the eventual development of the Poisson process, which is the paramount point process. Two notable Poisson applications, before the modern era of probability, were Erlang's (1909) model of telephone calls to a trunk line

and Bateman's (1910) model of α -particles emitted from a radioactive substance.

Life-Tables, Systems Reliability and Renewal Phenomena. The numerous studies of mortality based on life-tables from Graunt (1662) to Lotka (1939), and related studies in this century, such as Weibull's (1939) study of system lifetimes, were the precursors of renewal processes.

Queueing in Telecommunications. Palm's (1943) pioneering work on queueing in telecommunications and Khinchine's (1960) mathematical foundations of queueing processes showed the significance of modeling the flow of customers into a service station as a point process. This highlighted the need to develop point processes of event occurrences other than Poisson or renewal processes.

Statistical Mechanics. Gibbs' (1902) fundamental work on statistical mechanics was a major catalyst for developing point processes in spaces other than the real line and with interactions among the points.

Most of the theory of point processes has been developed in the last 30 years. The standard families of point processes are: (1) Poisson, compound Poisson and Cox processes. (2) Infinitely divisible and independent increment point processes. (3) Renewal processes and processes defined by interval properties. (4) Stationary point processes. (5) Marked point processes, which are associated with each of the other families. (6) Point processes related to Martingale theory and stochastic calculus.

In this chapter, I shall describe the structure of most of these processes and discuss some of their basic properties. The coverage does not include several important topics requiring lengthy mathematical development (e.g. martingale theory of point processes, general Palm probabilities, and

ergodic and spectral analysis of stationary processes). The emphasis will be on presenting tools for modeling stochastic systems rather than on applications of the tools. Although the theory of point processes is intimately connected with the subject of measure and integration (a point process is a random counting measure), I shall focus on results that can be understood without a deep knowledge of measure theory. On the other hand, the presentation will be rigorous and at the level of the applied probability literature that one would encounter in studying point processes.

The standard books and survey articles on point processes are as follows. Introductory works with a fairly broad coverage are Daley and Vere-Jones (1972) and Cox and Isham (1980). One can obtain a good introduction to point processes by reading these along with Cox and Lewis (1966), Grandell (1976, 1977), Snyder (1975), Jagers (1972), Karr (1986, Chapters 1, 2) and selected articles in Lewis (1972). The forthcoming book by Daley and Vere-Jones (1988) gives a detailed introduction and more comprehensive development of the theory. The most recent and rather complete research monograph on the mathematical theory of point processes and random measures is Kallenberg (1983).

In addition, there are a number of books on special topics. Feller (1971), Çinlar (1975), Gut and Prabhu (1987) and standard introductory texts on stochastic processes provide a good coverage of renewal processes. Khintchine (1960), Cramér and Leadbetter (1967), Franken, König, et al. (1981), Rolski (1981), Baccelli and Brémaud (1986), and Neveu (1977) are studies of stationary point processes and queueing (also see Bartfai and Tomkó (1981)). Matthes, Kerstan and Mecke (1978) discuss infinitely

divisible point processes. The martingale approach to point processes appears in Brémaud (1981), Liptser and Shirayev (1978), Ikeda and Watanabe (1981), and Elliott (1982). Also, Ripley (1981), Jacobsen (1982) and Karr (1986) study statistical inference and prediction problems of point processes.

Other subjects related to point processes are random sets (Kendall (1974), Matheron (1975) and Ripley (1976)), systems of interacting particles (Liggett (1985)), random fields (Kinderman and Snell (1980) and Rozanov (1980)), percolation processes (Kesten (1982)), and extreme value theory (Leadbetter, Lindgren and Rootzén (1983)).

To see what topics lie ahead, consult the table of contents. Most of the results herein are proved and developed further in Cox and Isham (1980), Daley and Vere-Jones (1988), Feller (1971), and Kallenberg (1983). The other references I cite are for particular points. For a complete set of references and a chronology of the development of point processes, see these books and also Karr (1986).

1.2 Definition of a Point Process

The classical definition of a point process on $R_+ \equiv [0, \infty)$ is as follows. Suppose that T_1, T_2, \dots are random variables on a probability space representing locations of points on R_+ , such as the successive times at which an event occurs. Assume that $0 \leq T_1 \leq T_2 \leq \dots$ and $T_n \rightarrow \infty$ a.s. (almost surely). Then the number of points or event occurrences in the interval $[0, t]$ is given by

$$N_t = \sum_{n=1}^{\infty} 1(T_n \leq t),$$

where $1(A)$ is the indicator of A (it is one on A and zero elsewhere). The

counting process $\{N_t; t \geq 0\}$, or the sequence $\{T_n; n=1,2,\dots\}$, is called a point process on R_+ .

This definition is adequate for many applications. However, a more general definition is required for (i) modeling point processes with a finite, random number of points (the process above has an infinite number), (ii) defining point processes on general spaces, and (iii) characterizing the probability law of a counting process directly without reference to its point locations. The characterization in (iii) is needed for analyzing sums (superpositions) and other operations on processes, studying the convergence of processes, comparing processes via order relations, and even constructing approximations.

We shall adopt the following definition of a point process. We let E denote the space in which the points lie. For our purposes, we assume that E is an Euclidean space (e.g. R_+ , R^d , $[a,b]$, a countable set), or a product of these (e.g. $R_+ \times C$, $R^d \times C$). The E could also be a more general topological space, see Kallenberg (1983). We distinguish three kinds of subsets of E : the class \mathcal{J} of all intervals or rectangles in E of the form $(a,b] = \{x \in E: a < x \leq b \text{ coordinatewise}\}$, the class \mathcal{E} of Borel sets of E (those formed from countable unions and intersections of sets in \mathcal{J}), and the class \mathcal{B} of bounded Borel sets (a set is bounded if it is contained in a finite interval). As usual, all functions herein are assumed to be measurable ($f: E \rightarrow R$ is measurable if $\{x: f(x) \leq a\} \in \mathcal{E}$ for each $a \in R$).

Definition 1.1. A point process N on E is a collection of non-negative integer-valued random variables $N = \{N(A): A \in \mathcal{E}\}$ on a probability space that take values in $\{0,1,\dots,\infty\}$ and satisfy the following conditions:

- (i) $N(\emptyset) = 0$ and $N(B) < \infty$ a.s. for each $B \in \mathcal{B}$.
- (ii) $N(\bigcup_n B_n) = \sum_n N(B_n)$ a.s. for any disjoint B_1, B_2, \dots in \mathcal{B} .

The random variable $N(A)$ represents the number of points in the set A .

A more formal but equivalent definition is as follows. A counting measure on E is a mapping μ from \mathcal{E} to $\{0, 1, \dots, \infty\}$ such that $\mu(\emptyset) = 0$; $\mu(B) < \infty$, $B \in \mathcal{B}$; and $\mu(\bigcup_n B_n) = \sum_n \mu(B_n)$ for disjoint B_1, B_2, \dots in \mathcal{B} . Let \mathcal{N} denote the set of all such counting measures. Note that the defining conditions (i), (ii) above for N simply say that almost every realization of N is a counting measure on E (i.e. an element of \mathcal{N}). This leads to the following definition.

Definition 1.1. A point process on E is a measurable mapping N from a probability space to \mathcal{N} .

We shall also use non-integer valued random measures. A random measure Λ on E is a collection $\Lambda = \{\Lambda(A) : A \in \mathcal{E}\}$ of non-negative random variables satisfying conditions (i) and (ii).

Suppose that N is a point process on E . Then there exist random variables X_1, X_2, \dots with values in E and a random variable ν with values in $\{0, 1, \dots, \infty\}$ such that

$$N(A) = \sum_{n=1}^{\nu} \delta_{X_n}(A), \quad A \in \mathcal{E},$$

where $\delta_x(B) = 1(x \in B)$ is the Dirac measure with unit mass at x . The X_n 's represent the locations of the points and $\nu = N(E)$ is the total number of points. We denote this representation by $N = \sum_{n=1}^{\nu} \delta_{X_n}$. We also let $N(a, b]$ denote $N((a, b])$.

Note that there are at most a finite number of X_n 's in a bounded set, and the subscripts on the X_n 's may not be unique. Also, N may have several points at one location. When $N(\{x\}) = 0$ a.s. for each $x \in E$, then we say N is simple (a sufficient condition for this is in Lemma 1.7). If N is not simple, then we can write

$$N(A) = \sum_{n=1}^{v'} Z_n \delta_{X'_n}(A)$$

where X'_1, X'_2, \dots are distinct point locations, $Z_n = N(\{X'_n\})$ is the number of points at X'_n and v' is the number of point locations.

When N is a point process on $E=R_+$, it is standard to use T_n 's instead of X_n 's and subscript them such that $0 \leq T_0 \leq T_1 \leq \dots$ a.s. Such a process is often used to model the occurrences of an event with T_n being the time of the n -th occurrence. Because of the total ordering of R_+ , and hence of the T_n 's, the theory of N is equivalent to that of the increasing stochastic process

$$N_t \equiv N[0, t], \quad t \geq 0.$$

(i.e. each $N(B)$ can be expressed in terms of N_t 's). Similarly, for N on $E=R$, the standard representation is

$$\dots T_{-2} \leq T_{-1} \leq T_0 \leq 0 < T_1 \leq T_2 \leq \dots$$

$$N(A) = \sum_{n=-v_1}^{v_2} \delta_{T_n}(A), \quad A \in \mathcal{E},$$

and the associated increasing process is

$$N_t \equiv \begin{cases} N(0, t] & t \geq 0 \\ -N(t, 0] & t < 0. \end{cases}$$

For $E = R^d$, there is no standard ordering of the X_n 's, but the last increasing process is still well defined (here $t \in R^d$ and $(a, b] \in \mathcal{I}$).

The moments of a point process N on E are as follows (the expectations below may be $+\infty$ except when $\infty - \infty$ is encountered):

The mean measure (or intensity measure) of N is $\mu(A) \equiv EN(A)$, $A \in \mathcal{E}$. These measures also arise in expectations of integrals with respect to N (Remark 1.5).

The k -th moment measure of N is

$$\mu_k(A_1 \times \dots \times A_k) \equiv E[N(A_1) \dots N(A_k)], \quad A_1, \dots, A_k \text{ in } \mathcal{E}$$

The covariance measure of N is

$$\text{Cov}(N(A), N(B)) \equiv \mu_2(A \times B) - \mu(A)\mu(B), \quad A, B \in \mathcal{E}.$$

The k-th factorial moment measure of N is

$$\mu_{(k)}(A) \equiv E[N(A)^{(k)}], \quad A \in \mathcal{E}, \text{ where } n^{(k)} = n(n-1)\cdots(n-k), \text{ and}$$

$$\mu_{(k)}(A_1^{k_1} \times \cdots \times A_r^{k_r}) \equiv E\left[N(A_1)^{(k_1)} \cdots N(A_r)^{(k_r)}\right], \quad A_1, \dots, A_r \text{ disjoint in } \mathcal{E}$$

and $k_1 + \dots + k_r = k$. This uniquely defines $\mu_{(k)}(B)$ for each $B \in \mathcal{E}^k$. The difference between $\mu_{(k)}$ and μ_k is easy to see for $k=2$:

$$\mu_{(2)}(A \times B) = \mu_2(A \times B) - \mu(A \cap B), \quad A, B \in \mathcal{E}.$$

Also, if μ_k has a density $f_k: E^k \rightarrow R_+$ so that

$$\mu_k(A) = \int_A f_k(x_1, \dots, x_k) dx_1 \dots dx_k,$$

then, in comparison,

$$\mu_{(k)}(A) = \int_{A'} f_k(x_1, \dots, x_k) dx_1 \dots dx_k,$$

where $A' = \{(x_1, \dots, x_k) \in A: x_1, \dots, x_k \text{ are distinct}\}$. The moment measures yield infinitesimal probabilities in the usual manner: When μ_k has the density f_k , then for distinct x_1, \dots, x_k ,

$$P(N(x_1, x_1+dx_1] = 1, \dots, N(x_k, x_k+dx_k] = 1) = f_k(x_1, \dots, x_k) dx_1 \dots dx_k.$$

1.3 Distributions and Laplace Functionals of Point Processes

The distribution or probability law of a point process N on E is uniquely determined by the joint distribution of $N(B_1), \dots, N(B_n)$, for each n and disjoint B_1, \dots, B_n in \mathcal{B} (called the finite dimensional distributions of N). The following is an elaboration on this. We use $\stackrel{d}{=}$ to mean equal in distribution.

Remarks 1.2. Let $N = \sum_{n=1}^{\infty} \delta_{X_n}$ and $N' = \sum_{n=1}^{\infty} \delta_{X'_n}$ be point processes on E (they may be defined on separate probability spaces).

(a) $N \stackrel{d}{=} N'$ if and only if $(N(I_1), \dots, N(I_n)) \stackrel{d}{=} (N'(I_1), \dots, N'(I_n))$ for each n and disjoint I_1, \dots, I_n in \mathcal{I} . When $E \subset \mathbb{R}^d$, this condition is equivalent (in terms of increasing processes) to $(N_{t_1}, \dots, N_{t_n}) \stackrel{d}{=} (N'_{t_1}, \dots, N'_{t_n})$, $t_1 \leq \dots \leq t_n$.

(b) When N and N' are simple, then $N \stackrel{d}{=} N'$ if and only if $P(N(B) = 0) = P(N'(B) = 0)$, $B \in \mathcal{B}$.

(c) If $(\nu, X_1, X_2, \dots) \stackrel{d}{=} (\nu', X'_1, X'_2, \dots)$, then $N \stackrel{d}{=} N'$. The converse is true when $E = \mathbb{R}_+$ or \mathbb{R} and the X_n 's are the ordered T_n 's.

The distribution of a point process is also uniquely determined by its Laplace functional as follows. We shall denote Lebesgue integrals as $\int_E f(x) \mu(dx)$, as opposed to $\int_E f(x) d\mu(x)$, and we sometimes omit the E . Readers unfamiliar with these integrals can interpret them as Riemann integrals $\int f(x) \phi(x) dx$, where ϕ is the density of μ (symbolically $\mu(dx) = \phi(x) dx$). These integrals also make sense when f or μ are random. An example is

$$\int_E f(x) N(dx) = \sum_{n=1}^{\nu} f(X_n), \quad \text{when } N = \sum_{n=1}^{\nu} \delta_{X_n}.$$

Definition 1.3. The Laplace functional of a point process N on E is

$$L_N(f) \equiv E\{\exp[-\int_E f(x) N(dx)]\}, \quad \text{where } f: E \rightarrow \mathbb{R}_+.$$

This is analogous to a Laplace transform $E(e^{-tZ})$ of a non-negative random variable Z . We first note that the joint Laplace transform of $N(A_1), \dots, N(A_n)$, for A_1, \dots, A_n in \mathcal{E} , is contained in L_N . Indeed, consider

$$(1.1) \quad \text{the simple function } f(x) = \sum_{k=1}^n t_k 1(x \in A_k). \quad \text{Then} \\ E\left\{\exp\left[-\sum_{k=1}^n t_k N(A_k)\right]\right\} = L_N(f).$$

Thus, L_N uniquely determines the joint distribution of $N(A_1), \dots, N(A_n)$, and

so L_N also uniquely determines the distribution of N . This, and a characterization of N via integrals are the subject of the next result. We let C_K denote the set of functions $f: E \rightarrow R_+$ that are continuous and such that $\{x: f(x) > 0\}$ is a bounded set.

Theorem 1.4. Suppose N and N' are point processes on E . The following statements are equivalent: (i) $N \stackrel{d}{=} N'$. (ii) $L_N(f) = L_{N'}(f)$, $f \in C_K$.

(iii) $\int f(x)N(dx) \stackrel{d}{=} \int f(x)N'(dx)$, $f \in C_K$.

Remark 1.5. Integrals of the form $\int_A f(x)N(dx)$, where $A \in \mathcal{E}$ and $f: E \rightarrow R$, are important in applications as well as in theoretical statements such as Theorem 1.4. A frequently used formula is

$$E\left[\int_A f(x)N(dx)\right] = \int_A f(x)\mu(dx),$$

where μ is the mean measure of N , provided the integral exists. This is proved by verifying it for simple functions and then limits of simple functions (as in the proof of Proposition 1.9).

The uses of Laplace functionals are similar to those of Laplace transforms. For example, moments of $N(A_1), \dots, N(A_k)$ can be obtained from derivatives of (1.1). As another example, suppose N_1, \dots, N_k are point processes on E . Then their sum or superposition $N = N_1 + \dots + N_k$ is also a point process (regardless of the dependency among the N_k 's). Now, if N_1, \dots, N_k are independent, then the Laplace transform of N has the product form

$$L_N(f) = E\left\{\exp\left[-\sum_{k=1}^n \int f(x)N_k(dx)\right]\right\} = \prod_{k=1}^n L_{N_k}(f),$$

which is sometimes convenient for obtaining the distribution of N .

1.4 Basic Examples: Poisson, Renewal and Stationary Processes

Point processes are commonly classified by their distributions (e.g. a compound Poisson process) or by a certain characteristic (e.g. a stationary

point process). They may also be classified by how they arise as a function of another process (e.g. the times when customers exit a network of service systems), by the mathematical techniques involved in their analysis, or by their application context. We now introduce several basic processes.

Suppose N is a point process on E . We say that N has independent increments if $N(B_1), \dots, N(B_n)$ are independent for any disjoint B_1, \dots, B_n in \mathcal{A} (we discuss these processes in § 3.3). We say that N is a stationary point process on $E \subset \mathbb{R}^d$ (or has stationary increments) if, for each B_1, \dots, B_n in \mathcal{A} ,

$$(N(B_1+x), \dots, N(B_n+x)) \stackrel{d}{=} (N(B_1), \dots, N(B_n)), \quad x \in E.$$

Here $B + x = \{y + x : y \in B\}$. Stationary processes are the subject of § 5.

The most important point process is the Poisson process.

Definition 1.6. A point process N on E is a Poisson process with mean measure Λ if N has independent increments and, for each $B \in \mathcal{A}$,

$$P\{N(B) = n\} = \Lambda(B)^n e^{-\Lambda(B)} / n!, \quad n=0,1,\dots$$

(This probability statement also holds for each B in the larger class \mathcal{E} , with $N(B) = 0$ or ∞ a.s. when $\Lambda(B) = 0$ or ∞ , respectively.) We call N a stationary Poisson process with rate λ when $\Lambda(dx) = \lambda dx$ for some $\lambda > 0$ and $E \subset \mathbb{R}^d$ (the general process is sometimes called non-stationary or non-homogeneous).

A common assumption for a Poisson process N (or any point process) is that its mean Λ takes the form $\Lambda(A) = \int_A \lambda_x dx$, or $\Lambda(dx) = \lambda_x dx$; this λ is sometimes called the rate or intensity of N . Note that N is simple if and only if $\Lambda(\{x\}) = 0$ (which is true when Λ has the preceding form). In general, when $\Lambda(\{x\}) > 0$, then $N(\{x\})$ is Poisson with that mean.

The prominence of Weiner and more general Gaussian processes is due primarily to the central limit phenomenon that these processes arise as limits of processes of sums of random variables (see for instance Billingsley

(1968)). Similarly, the prominence of Poisson processes stems from the property that they arise as limits of sums of uniformly sparse point processes (see Theorem 3.4). They also arise as limits of processes of rare events (see Theorem 3.7). Many actual point processes can indeed be viewed as superpositions of points from many sources or as processes of rare events. In addition, Poisson processes are building blocks for more complicated processes.

There are a number of characterizations of Poisson processes. The following one is based on the null probabilities $P\{N(A) = 0\}$. We first present a sufficient condition for simplicity (see p. 203 of Jagers (1972) for the proof).

Lemma 1.7. Suppose N is a point process on E and there is a measure Λ on E such that $\Lambda(\{x\}) = 0$, $x \in E$ and $P\{N(B) \geq 2\} = o(\Lambda(B))$ as $\Lambda(B) \rightarrow 0$. Then N is simple.

Proposition 1.8. Suppose N is a point process on E that satisfies the hypothesis of Lemma 1.7 and $P\{N(B) = 0\} = e^{-\Lambda(B)}$, $B \in \mathcal{B}$. Then N is a Poisson process with mean measure Λ .

Proof. By Lemma 1.7, N is simple and so by Remark 1.2(b), the distribution of N is uniquely determined by the null probabilities $P\{N(B) = 0\}$. But these, by assumption, are those of a Poisson process with mean Λ .

The next result gives an expression for the Laplace functional of Poisson processes. The proof demonstrates a common approach for deriving Laplace functionals in general. We shall return to Poisson processes in the next section.

Proposition 1.9. If N is a Poisson process with mean measure Λ then

$$L_N(f) = \exp[-\int (1 - e^{-f(x)}) \Lambda(dx)].$$

Proof. Denote the right side by $G(f)$. First, consider the simple function

$f(x) = \sum_{k=1}^n t_k 1(x \in B_k)$, where B_1, \dots, B_n are disjoint in \mathfrak{B} . Since

$N(B_1), \dots, N(B_n)$ are independent Poisson variables,

$$\begin{aligned} L_N(f) &= E\{\exp[-\sum_{k=1}^n t_k N(B_k)]\} = \prod_{k=1}^n E[e^{-t_k N(B_k)}] \\ &= \prod_{k=1}^n \exp[-\lambda(B_k)(1 - e^{-t_k})] = G(f). \end{aligned}$$

Next, consider a function f such that $\{x: f(x) > 0\} \in \mathfrak{B}$. We can write $f(x) = \lim_{n \rightarrow \infty} f_n(x)$, $x \in E$, where the f_n 's are simple functions as above. Then by two applications of the dominated convergence theorem and the preceding result, we have

$$L_N(f) = \lim_{n \rightarrow \infty} L_N(f_n) = \lim_{n \rightarrow \infty} G(f_n) = G(f).$$

Finally, consider any function f , and let B_n be a sequence in \mathfrak{B} such that $B_n \uparrow E$. Define $f_n(x) \equiv f(x) 1(x \in B_n)$, and so $\{x: f_n(x) > 0\} \in \mathfrak{B}$. Then the preceding equalities hold by two applications of the monotone convergence theorem.

For point processes on R_+ or R , it is common to specify their distributions by specifying the distributions of their point locations or their inter-point distances. The primary example is the renewal process.

Definition 1.10. Let $N = \sum_{n=1}^{\infty} \delta_{T_n}$ be a point process on R_+ , and let $W_1 = T_1$ and $W_n = T_n - T_{n-1}$, $n=2,3,\dots$. The N is a renewal process with waiting time distribution F if W_1, W_2, \dots are independent and each one has the distribution F . For simplicity, we assume that $F(0) = 0$, and so N is simple. We call $T_1 < T_2 < \dots$ the renewal times of N and W_1, W_2, \dots the waiting times between renewals.

Remark 1.11. The distribution of a renewal process is uniquely determined by its waiting time distribution. To see this, suppose N and N' are two renewal processes. By Remark 1.2(c) we know that $N \stackrel{d}{=} N'$ if and only if $(T_1, T_2, \dots) \stackrel{d}{=} (T'_1, T'_2, \dots)$. But the latter is equivalent to $(W_1, W_2, \dots) \stackrel{d}{=} (W'_1, W'_2, \dots)$ which is equivalent to $F = F'$. Thus $N \stackrel{d}{=} N'$ if and only if $F = F'$. This uniqueness property is used, for instance, in stochastic comparisons or convergence results.

Although renewal processes have a simple structure, their Laplace functionals generally do not. We shall continue our discussion of renewal processes in §4.

1.5 Marked and Compound Point Processes

In addition to their locations, the points of a point process may have distinguishing attributes or attendant information, which are commonly called marks. The standard way of modeling marks is as follows.

Consider a point process $N = \sum_{n=1}^{\nu} \delta_{X_n}$ on E . Suppose that associated with the point at X_n ($n \leq \nu$) there is a mark Z_n that takes values in some space E' . Then the point process M on $E \times E'$ defined by

$$M(A \times A') = \sum_{n=1}^{\nu} 1(X_n \in A, Z_n \in A'), \quad A \times A' \in \mathcal{E} \times \mathcal{E}'$$

represents the marks as well as the point locations.

Before we formalize this notion of marked point processes, consider the special case in which the marks Z_n are conditionally independent given N and a mark for a point at x has the conditional distribution $K(x, dz)$ (often called a kernel); that is, for A_1, \dots, A_n in \mathcal{E}' ,

$$P\{Z_1 \in A_1, \dots, Z_n \in A_n \mid N, \nu \geq n\} = \prod_{j=1}^n K(X_j, A_j).$$

In other words, the mark Z_n at $X_n = x$, when $v \geq n$, depends on N only through x . Then the Laplace functional of M is as follows.

Lemma 1.12. For $f: E \times E' \rightarrow \mathbb{R}_+$,

$$L_M(f) = L_N(g).$$

where $g(x) = -\log \int_{E'} e^{-f(x,z)} K(x,dz)$.

Proof. Since the Z_n 's are conditionally independent given N , therefore

$$\begin{aligned} L_M(f) &= E\{E[\exp(-\sum_{n=1}^v f(X_n, Z_n)) | N]\} \\ &= E\left\{\prod_{n=1}^v E\left[e^{-f(X_n, Z_n)} \mid N, v \geq n\right]\right\} \\ &= E\left\{\exp\left[-\sum_{n=1}^v g(X_n)\right]\right\} = L_N(g). \end{aligned}$$

We are now ready for our definition.

Definition 1.13. Suppose N is a point process on E and M is a point process on $E \times E'$ such that $N(A) = M(A \times E')$, $A \in \mathcal{E}$. We call M a marked point process of N . Furthermore, we say that M has location-dependent marks with distribution $K(x,dz)$ if the Laplace functional of M is as in Lemma 1.12.

Keep in mind that a typical representation of a marked point process is

$$M = \sum_{n=1}^v \delta_{X_n, Z_n} \text{ where } N = \sum_{n=1}^v \delta_{X_n}. \text{ If } M \text{ has location-dependent marks with}$$

distribution $K(x,dz)$, then the mean measure of M is (recall Remark 1.5),

$$E[M(A \times A')] = E\{E[M(A \times A') | N]\} = E\left\{\int_A K(x, A') N(dx)\right\} = \int_A K(x, A') \mu(dx),$$

where μ is the mean measure of N . That is, $EM(dxdz) = K(x,dz)\mu(dx)$.

Moreover, for $f: E \times E' \rightarrow \mathbb{R}$,

$$E\left[\int_A \int_{A'} f(x,z) M(dxdz)\right] = \int_A \int_{A'} f(x,z) K(x,dz) \mu(dx).$$

If N is a point process on R_+ with marks that are real-valued (or are vectors or elements of a semigroup), then these marks are often modeled by the cumulative process

$$Z(t) = \sum_{n=1}^D Z_n 1(T_n \leq t), \quad t \in R_+.$$

Alternate representations are

$$Z(t) = \int_0^t \int_{E'} z M(dx dz) = \sum_{n=1}^{N_t} Z_n.$$

This process provides the same information as the marked process M .

Example 1.14 Marks as Functionals of a Stochastic Process. Suppose $Y = \{Y_t; t \geq 0\}$ is a real-valued regenerative process over the times $0 < T_1 < T_2 < \dots$ (see § 4.1). A typical mark of T_n might be the discounted cost

$$Z_n = \int_{T_{n-1}}^{T_n} e^{-\alpha t} c(Y_t) dt,$$

where $c(y)$ is the cost per unit time of Y being in state y . Then the cumulative process

$$Z(t) = \sum_{n=1}^{N_t} Z_n = \int_0^{T_{N_t}} e^{-\alpha s} c(Y_s) ds$$

is the discounted cost in the interval $[0, T_{N_t}]$. This and a variety of other marks are functionals of the form

$$Z_n = \phi(T_{n-1}, T_n, \{Y_t; t \in (T_{n-1}, T_n]\}).$$

Another example is

$$Z_n = \max\{Y_t; t \in (T_{n-1}, T_n]\}.$$

In situations where points occur in batches, an appropriate model might be a compound point process defined as follows.

Definition 1.15. Suppose N is a point process on E and M is a marked point process of N on $E \times \{0, 1, \dots\}$ with location-dependent marks having distribution $K(x, dz)$. The point process

$$N'(A) \equiv \int_A \int_{E'} z M(dx dz) = \sum_{n=1}^{\infty} Z_n \delta_{X_n}(A), \quad A \in \mathcal{E},$$

is a compound point process of N with mass distribution $K(x, dz)$. The term compound has traditionally been used only for location-independent marks where the Z_n 's are independent of N . For instance, when N is a Poisson process with mean Λ and $K(x, dz) = F(dz)$ independent of x , then N' is a compound Poisson process whose points have the mean measure Λ and mass distribution F . (This N' is a compound Poisson random measure when the Z_n 's are real-valued instead of integer-valued.)

Similar to Lemma 1.12, the Laplace functional of the compound point process N' is

$$L_{N'}(f) = L_N(h), \quad \text{where } h(x) = - \log \sum_{z=0}^{\infty} e^{-zf(x)} K(x, \{z\}),$$

and that for the compound Poisson process N' is

$$(1.2) \quad L_{N'}(f) = \exp \left[- \int_E \sum_{z=0}^{\infty} (1 - e^{-zf(x)}) \Lambda(dx) F(\{z\}) \right].$$

2. POISSON PROCESSES AND SOME RELATIVES

In this section, we characterize the structure of Poisson processes and discuss several operations on them including sums, partitions, thinnings and translations. We also give brief descriptions of sample processes, Cox processes, negative binomial processes and cluster processes.

2.1 Characterization of Poisson Processes

The following results show how the distributions of the counting variables for a Poisson process determine the distributions of its point locations. Stationary Poisson processes on R_+ or R are especially nice.

Theorem 2.1. Suppose N is a stationary Poisson process on R_+ with rate λ . Then the interpoint distances W_1, W_2, \dots are independent exponentially distributed with mean λ^{-1} and the location of the n th point $T_n = \sum_{k=1}^n W_k$ has a gamma distribution with order n and scale parameter λ .

Proof. By the definition of a Poisson process, it follows that, for each n and $t > 0$,

$$\begin{aligned} P\{W_{n+1} > t | W_1, \dots, W_n\} &= P\{N(T_n, T_n+t] = 0 | T_n, N(B), BC(0, T_n]\} \\ &= P\{N[(0, t] = 0\} = e^{-\lambda t}. \end{aligned}$$

Using this in an induction argument proves the assertion about the W_n 's. The gamma distribution of T_n is a standard consequence.

The preceding result and Remark 1.12(c) yield the following characterization: A renewal process is a Poisson process if and only if its waiting time distribution is exponential.

The next result characterizes the point locations of Poisson processes on general spaces.

Theorem 2.2. A point process N on E is a Poisson process with mean measure Λ if and only if, for each $B \in \mathcal{B}$, the process N on B is equal in distribution to a point process $N' = \sum_{n=1}^v \delta_{U_n}$ on B , where v, U_1, U_2, \dots are independent random variables such that v has a Poisson distribution with mean $\Lambda(B)$ and each U_n takes values in B and has the distribution

$$F(A) = \Lambda(A)/\Lambda(B), \quad ACB.$$

Proof. It suffices, by Theorem 1.4, to show that N' has the Poisson Laplace functional shown in Proposition 1.9. But this follows since, by the properties of v, U_1, U_2, \dots , we have

$$\begin{aligned} L_{N'}(f) &= E(E\{\exp[-\sum_{n=1}^v f(U_n)] | v\}) = E([\int e^{-f(x)} F(dx)]^v) \\ &= \exp\{-\Lambda(B)[1 - \int e^{-f(x)} F(dx)]\} = \exp[-\int (1 - e^{-f(x)}) \Lambda(dx)]. \end{aligned}$$

This result says that on a bounded set B , the Poisson number of points $N(B)$ are located as a random sample U_1, U_2, \dots from the distribution F . This characterization is useful for deciding whether a Poisson process is an appropriate model for a certain phenomenon. It is also useful for

applications; note, for instance, that $\int_B f(x) N(dx) \stackrel{d}{=} \sum_{n=1}^v f(U_n)$.

Remark 2.3. For a point process on R_+ , the preceding result is usually expressed as follows. Suppose that N is a Poisson process on R_+ with mean Λ and point locations $0 \leq T_1 \leq T_2 \leq \dots$. Then, for a given $t > 0$, the conditional distribution of T_1, \dots, T_n given $N_t = n$ is equal to the distribution of the order statistics $U_{(1)} \leq U_{(2)} \leq \dots \leq U_{(n)}$ of U_1, \dots, U_n that are independent with distribution $F(s) = \Lambda_s/\Lambda_t$, $0 \leq s \leq t$. That is, for $0 < t_1 < u_1 < \dots < t_n < u_n < t$,

$$\begin{aligned}
 (2.1) \quad & P\{T_1 \in (t_1, u_1], \dots, T_n \in (t_n, u_n] | N_t = n\} \\
 & = n! (\Lambda_{t_1} - \Lambda_{u_1}) \dots (\Lambda_{t_n} - \Lambda_{u_n}).
 \end{aligned}$$

When $\Lambda_t = \int_0^t \lambda_s ds$, this distribution has the density

$$f_{T_1, \dots, T_n | N_t = n}(t_1, \dots, t_n) = n! \lambda_{t_1} \dots \lambda_{t_n} / \Lambda(t)$$

For the special case when N is stationary with rate λ , then U_1, \dots, U_n are uniformly distributed on $[0, t]$, and the last conditional density reduces to $n!/t^n$.

Another immediate consequence of Theorem 2.2 is as follows. Suppose N is a Poisson process on E with mean Λ . Then for each n , $B \in \mathcal{B}$, disjoint A_1, \dots, A_k in \mathcal{A} and $n_1 + \dots + n_k = n$,

$$(2.2) \quad P\{N(A_1) = n_1, \dots, N(A_k) = n_k | N(B) = n\} = \frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k}$$

where $p_i = \Lambda(A_i)/\Lambda(B)$. This is the multinomial probability that n_1, \dots, n_k points in the sample U_1, \dots, U_n fall in the respective sets A_1, \dots, A_k where $p_i = P(U_m \in A_i)$.

2.2 Sample Processes

The order statistic property of Poisson processes in Theorem 2.2 is characteristic of the following important family of finite point processes.

Suppose X_1, X_2, \dots are independent random variables that take values in E and each one has the distribution F . Then $N = \sum_{k=1}^n \delta_{X_k}$ is a point process on

E with

$$L_N(f) = E\left\{\exp\left[-\sum_{k=1}^n f(X_k)\right]\right\} = \left[\int_E e^{-f(x)} F(dx)\right]^n.$$

Any point process on E that is equal in distribution to N (or has the preceding Laplace functional) is called a sample process on E with mean measure nF . Now, suppose ν is a non-negative integer-valued random variable independent of N . Then $N' = \sum_{k=1}^{\nu} \delta_{X_k}$ is a point process on E with

$$L_{N'}(f) = G\left(\int_E e^{-f(x)} F(dx)\right), \text{ where } G(s) = E(s^{\nu}).$$

Any point process on E that is equal in distribution to N' is called a mixed sample process with mean measure $E\nu F$ (N' is N with ν randomized and hence is a mixture of N).

Example 2.4. A Dispatching Model. Suppose that a random number ν of items (e.g. parts or data packets) arrive at a station in a fixed time interval $[0, T]$. The times at which the items arrive are independent with distribution F , and are independent of ν . At fixed times $0 \equiv t_0 < t_1 < \dots < t_n \equiv T$, the accumulated items are instantaneously dispatched from the station. Then the total waiting time of the items at the station up to time T is

$$W(t_1, \dots, t_n) \equiv \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} N(t_k, t_k + u) du,$$

where N is the mixed sample process defined above. Note that

$$EW(t_1, \dots, t_n) = E\nu \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} [F(t_k + u) - F(t_k)] du$$

(the expectation can be taken inside the integral by Fubini's theorem). This expression can be used to optimize the dispatch times t_1, \dots, t_n and the number n of dispatches as well. For instance, if there is a cost c for each dispatch and a cost h per unit time for holding each item, then the problem is

$$\min_{n, t_1, \dots, t_n} \{cn + hE[W(t_1, \dots, t_n)]\}.$$

For the case in which N is a stationary Poisson process with rate λ , one can

show by a calculus argument that the optimal n^* is one of the integers adjacent to $T(h\lambda/2c)^{1/2}$ and the optimal $t_k^* = KT/n^*$ (equally spaced times).

Example 2.5. Discounted Costs. Suppose $N = \sum_{n=1}^v \delta_{T_n}$ is a point process of times in a period $[0, T]$ at which a company incurs costs Z_1, \dots, Z_v . Then the discounted cost up to time t is

$$C_t = \sum_{n=1}^{N_t} Z_n e^{-\alpha T_n}, \quad 0 \leq t \leq T.$$

Suppose N is a mixed sample process as above and Z_1, Z_2, \dots are independent with distribution G and are independent of N . Let $M = \sum_{n=1}^v \delta_{T_n} \cdot Z_n$. Clearly,

$EM(dt dz) = E\nu F(dt) G(dz)$. Then

$$EC_t = E \left[\int_0^t \int_R z e^{-\alpha s} M(ds dz) \right]$$

$$= EZ_1 E\nu \int_0^t e^{-\alpha s} F(ds).$$

For the special case in which N is a stationary Poisson process with rate λ , so that $F(dt) = T^{-1} dt$ and $E\nu = \lambda T$, then $EC_t = (\lambda/\alpha) EZ_1 (1 - e^{-\alpha t})$.

2.3 Sums, Partitions, Thinnings and Translations of Poisson Processes

Some standard operations on point processes are summing of processes, partitioning a process into several subprocesses, thinning (deleting) points in a process, and translating the points in a process. We shall now discuss these operations for Poisson processes.

We first observe that a sum of independent Poisson processes is also Poisson.

Theorem 2.6. If N_1, \dots, N_n are independent Poisson processes on E with respective mean measures $\Lambda_1, \dots, \Lambda_n$, then $N = N_1 + \dots + N_n$ is a Poisson process with mean measure $\Lambda = \Lambda_1 + \dots + \Lambda_n$.

Proof. This follows immediately from the defining properties of a Poisson process. Another approach is to observe that

$$\begin{aligned} L_N(f) &= \prod_{k=1}^n L_{N_k}(f) = \prod_{k=1}^n \exp \left[- \int (1 - e^{-f(x)}) \Lambda_k(dx) \right] \\ &= \exp \left[- \int (1 - e^{-f(x)}) \Lambda(dx) \right]. \end{aligned}$$

This result extends to infinite sums $N = \sum_{k=0}^{\infty} N_k$ when $\Lambda(B) < \infty$ for $B \in \mathcal{B}$.

It also holds for other families of processes including stationary, Cox (§2.4), infinitely divisible and independent increment processes (§ 3.3), and mixed sample processes with a common sample distribution. On the other hand, it does not hold for renewal processes or stationary interval processes (§ 4.5).

A large class of marked point processes of Poisson processes are also Poisson.

Proposition 2.7. Suppose N is a Poisson process on E with mean Λ and M is a marked point process of N on $E \times E'$ with position-dependent marks having distribution $K(x, dz)$. Then M is a Poisson process with mean measure $\Lambda(dx)K(x, dz)$.

Proof. This follows since Proposition 1.9 and Lemma 1.12 yield

$$\begin{aligned} L_M(f) &= L_N(g) = \exp \left[- \int (1 - e^{-g(x)}) \Lambda(dx) \right] \\ &= \exp \left[- \int_{E \times E'} (1 - e^{-f(x, z)}) \Lambda(dx) K(x, dz) \right]. \end{aligned}$$

We now discuss partitions and thinnings of a point process N on E . By a partition of N we mean any collection of point processes N_1, \dots, N_n on E such that $N = N_1 + \dots + N_n$. Typically, N is a "parent" process and each of its points is assigned randomly, by some rule, to one of the subprocesses

N_1, \dots, N_n . There is a one-to-one correspondence between partitions N_1, \dots, N_n of N and marked point processes M of N on $E \times \{1, \dots, n\}$; namely,

$$N_k(A) \equiv M(A \times \{k\}) = \sum_{n=1}^v 1(X_n \in A, Z_n = k), \quad A \in E, k=1, \dots, n.$$

The mark Z_n indicates the subprocess to which the point at X_n is assigned. As an example, suppose $X_1 < X_2 < \dots$ are times at which data packets enter a computer data base consisting of n files and Z_1, Z_2, \dots are the respective files to which the packets are sent. Then N_k models the number of packets entering file k over time. Analogous partitions depict multiple customer flows in queueing networks, part flows in manufacturing systems, and demand occurrences in economic markets. In some cases the partitioning rule is implicit: If each point of N has one of n attributes, then N_1, \dots, N_n can serve as models of the numbers of points with these respective attributes.

The notion of thinning of the point process N refers to the operation of randomly deleting some of its points. In general, a point process N' on E such that $N'(A) \leq N(A)$, $A \in E$, is called a thinning of N (N' is a thinner version of N). Any subprocess in a partition of N is a thinning of N .

We shall consider the following basic partitions and thinnings. We say that N_1, \dots, N_n is a partition of N based on the probabilities $p_1(x), \dots, p_n(x)$ if a point of N at x is assigned to subprocess N_k with probability $p_k(x)$ independent of everything else (i.e. Z_1, Z_2, \dots are position-dependent with distribution $p_k(t) = P\{Z_n = k \mid X_n = x, n \geq t\}$). Similarly, we say that N' is a $p(x)$ - thinning of N if a point of N at x is retained (assigned to N') with probability $p(x)$ independent of everything else.

Theorem 2.8. Suppose N is a Poisson process on E with mean Λ . If N_1, \dots, N_n is a partition of N based on the probabilities $p_1(x), \dots, p_n(x)$, then N_1, \dots, N_n are independent Poisson processes with respective mean measures

$p_1(x)\lambda(dx), \dots, p_n(x)\lambda(dx)$. Hence, if N' is a $p(x)$ -thinning of N , then N' is a Poisson process with mean measure $p(x)\lambda(dx)$.

Proof. Let M be the marked point process on $E \times \{1, \dots, n\}$ such that $N_k(A) = M(A \times \{k\})$. Since N is Poisson, Proposition 2.7 implies that M is also Poisson with mean $EM(dx \times \{k\}) = p_k(x)\lambda(dx)$. Thus each N_k is Poisson with mean $p_k(x)\lambda(dx)$. Moreover, N_1, \dots, N_n are independent since they are the restriction of the Poisson process M to the disjoint sets $E \times \{1\}, \dots, E \times \{n\}$, respectively.

Example 2.9. Non-stationary Poisson Processes as Time Transformations or Thinnings of Stationary Processes. Let N be a Poisson process on R_+ with mean $\Lambda_t = \int_0^t \lambda_s ds$. Let N_1 be a stationary Poisson process on R_+ with rate 1, and let $N_1 \circ \Lambda$ denote the point process on R_+ defined by $N_1 \circ \Lambda(a, b] \equiv N_1(\Lambda_a, \Lambda_b]$, $a < b$. Clearly $N_1 \circ \Lambda$ has independent increments and $N_1 \circ \Lambda_t$ is Poisson with mean Λ_t . Therefore, $N \stackrel{d}{=} N_1 \circ \Lambda$. That is, N is a time transformation of the stationary Poisson process N_1 . Another characterization of N is as follows. Consider N on $E = [0, T]$ where $\Lambda_T < \infty$ and let N' be a $p(x)$ -thinning of N_1 with $p(x) = \lambda_x / \Lambda_T$, $0 \leq x \leq T$. then Theorem 2.8 yields $N \stackrel{d}{=} N'$. This representation is convenient for simulations of N (Lewis and Shedler (1979)). One need only generate points of N_1 on E , say T_1, \dots, T_n and then retain the point at T_j with probability $p(T_j)$. The resulting points constitute N . This procedure requires the evaluation of only $\lambda_{T_j} / \Lambda_T$, $j=1, \dots, n$, (the integral Λ_t need not be evaluated as in the time-transformation characterization). Note that this procedure can also be used for $E \subset R^d$.

Another common operation on a point process is a random translation of its points. Suppose N is a point process on E and E is closed under addition

$(x + z \in E \text{ for each } x, z \in E)$. Suppose N' is a point process on E of the form,

$$N'(A) \equiv \int_{E^2} 1(x+z \in A) M(dx dz) = \sum_{n=1}^v 1(X_n + Z_n \in A), \quad A \in E,$$

where $M = \sum_{n=1}^v \delta_{X_n, Z_n}$ is a marked point process of N on E^2 whose marks are independent with distribution F . That is, N' is the process N with each point at x translated a random distance with distribution F (i.e. X_n is translated to $X_n + Z_n$). From Lemma 1.13, we have

$$L_{N'}(f) = E \left\{ \exp \left[- \int_{E^2} f(x+z) M(dx dz) \right] \right\} = L_N(g)$$

where $g(x) = - \log \int_E e^{-f(x+z)} F(dz)$. Here is a special case.

Theorem 2.10. If N is a Poisson process with mean measure Λ , then its translated version N' defined above is a Poisson process with mean measure $\Lambda'(A) = \int_E \Lambda(A-z) F(dz)$.

Proof. This follows since a substitution of the Poisson Laplace functional L_N in the preceding expression yields

$$\begin{aligned} L_{N'}(f) &= \exp \left[- \int_{E^2} (1 - e^{-f(x+z)}) F(dz) \Lambda(dx) \right] \\ &= \exp \left[- \int_E (1 - e^{-f(y)}) \Lambda'(dy) \right]. \end{aligned}$$

2.4 Cox Processes

One can construct natural generalizations of random variables by randomizing their parameters. For example, a random variable with distribution $P(X \leq x) = \int_0^\infty (1 - e^{-\lambda x}) F(d\lambda)$ is called an F-mixture of exponential distributions or an exponential random variable with a random parameter that has the distribution F . This procedure of randomizing parameters also extends to stochastic processes (recall the mixed sample

process). Another primary example is the Cox process: a Poisson process with randomized intensity.

Definition 2.11. Let N be a point process on E and Λ be a random measure on E (not necessarily on the same probability space as N). The N is a Cox process directed by Λ if, for each n_1, \dots, n_k and disjoint A_1, \dots, A_k in E ,

$$P\{N(A_1) = n_1, \dots, N(A_k) = n_k\} = E \left[\prod_{j=1}^k \Lambda(A_j)^{n_j} e^{-\Lambda(A_j)} / n_j! \right].$$

This is equivalent to

$$L_N(f) = E \left\{ \exp \left[- \int_E (1 - e^{-f(x)}) \Lambda(dx) \right] \right\}.$$

Cox (1955) introduced this process, which is sometimes called a doubly stochastic Poisson process, a conditional Poisson process, or a Poisson process in a random environment.

As an example, suppose N is a point process on R_+ that represents the number of failures over time of a robot in a manufacturing plant. The robot makes several types of parts and when it is used on a job of type u , its number of failures is a Poisson process with constant rate $\alpha(u)$.

Consequently, if the production schedule were specified by a non-random function $u(t)$, then N would be a Poisson process with mean measure

$\Lambda_t = \int_0^t \alpha(u(s)) ds$. Suppose, however, that the production schedule is a stochastic process $\{U_t; t \geq 0\}$ that is not affected by the failures. Then the failure process N is a Cox process directed by Λ , where $\Lambda_t = \int_0^t \alpha(U_s) ds$, $t \geq 0$.

Since Cox processes are essentially Poisson processes, each result for Poisson processes generally has a counterpart for Cox processes. Some elementary properties of Cox processes are as follows (see Brémaud (1981), Grandell (1976), Karr (1986), and Synder (1975) for further discussion). Here N is a Cox process directed by Λ .

(a) $EN(A) = E\Lambda(A), \quad \text{Var}N(A) = E\Lambda(A) + \text{Var}\Lambda(A)$

$$M_{[k]}(\Lambda_1, \dots, \Lambda_k) = E[\Lambda_1(\Lambda_1) \dots \Lambda_k(\Lambda_k)].$$

(b) N is simple if and only if $\Lambda(\{x\}) = 0$ a.s., $x \in E$.

(c) N is stationary if and only if Λ is stationary (the definition being similar to that for point processes). When $\Lambda(A) = \int_A \lambda_x dx$, then Λ is stationary if and only if $\{\lambda_x: x \in E\}$ is a stationary process (cf. §5.1).

(d) When $E = \mathbb{R}_+$, then $N \stackrel{d}{=} N_1 \circ \Lambda$, recall Example 2.9, where N_1 is a stationary Poisson processes with rate 1 independent of Λ .

(e) The Poisson results in the preceding sections readily extend to N . For instance, if N_1, \dots, N_n is a partition of N based on the probabilities $p_1(x), \dots, p_n(x)$, then each N_k is a Cox process directed by $p_k(x)\Lambda(dx)$ and N_1, \dots, N_n are conditionally independent given Λ .

Example 2.12. Negative Binomial Processes. Let N be a Cox process on E directed by Λ . Suppose $\Lambda = Y\mu$, where μ is a non-random measure on E and Y is a random variable. Such a Cox process is sometimes called a mixed Poisson process. Then the Laplace functional of N reduces to

$$L_N(f) = \phi\left(\int_E (1 - e^{-f(x)}) \mu(dx)\right), \quad \text{where } \phi(s) = E(e^{-sY}).$$

Now consider the special case in which Y has a gamma distribution with $\phi(s) = (1+s)^{-r}$. Then

$$L_N(f) = \left[1 + \int_E (1 - e^{-f(x)}) \mu(dx)\right]^{-r}.$$

In particular, for $\Lambda_1, \dots, \Lambda_k$ in E ,

$$E\left\{\exp\left[-\sum_{k=1}^n t_k N(\Lambda_k)\right]\right\} = \left[p/(1 - \sum_{k=1}^n q_k t_k)\right]^r, \quad t_k \in [0,1],$$

where $p = \left[1 + \lambda\left(\bigcup_{i=1}^n \Lambda_i\right)\right]^{-1}$ and $q_k = p\lambda(\Lambda_k)$. This is the multivariate

Laplace transform of the n -dimensional negative binomial distribution with parameters n, p, q_1, \dots, q_n . That is, the finite dimensional distributions are

negative binomial. Accordingly, N is called a negative binomial point process (Gregoire (1983) and Diggle and Milne (1983)).*

The preceding is not the only type of negative binomial process. Here is another one. Let $N = \sum_{n=1}^{\infty} Z_n \delta_{T_n}$ be a compound Poisson process on E whose Poisson points have mean λ and whose masses Z_n have the probability density $P\{Z_n = m\} = (\rho^m/m) \log[1/(1-\rho)]$, $m=1,2,\dots$, where $0 < \rho < 1$. Then from (1.2), we have

$$\begin{aligned} L_N(f) &= \exp \left\{ - \sum_{m=1}^{\infty} \int_E (1 - e^{-mf(x)}) \lambda(dx) (\rho^m/m) \log[1/(1-\rho)] \right\} \\ &= \exp \int \log [(1 - \rho e^{-f(x)})/(1-\rho)] / \log(1-\rho) \lambda(dx). \end{aligned}$$

In particular, for disjoint A_1, \dots, A_n ,

$$E \left\{ \exp \left[- \sum_{k=1}^n t_k N(A_k) \right] \right\} = \prod_{k=1}^n [(1-\rho)/(1-\rho t_k)]^{-\lambda(A_k)/\log(1-\rho)}.$$

Again, this is an n -variate negative binomial Laplace transform. In this case, N has independent increments, which was not true for the preceding example.

2.5 Poisson Cluster Processes

Let $N = \sum_{n=1}^{\nu} \delta_{X_n}$ be a point process on E . Suppose there are point processes N_1, \dots, N_{ν} on some space E' representing clusters of points associated with the respective points X_1, \dots, X_{ν} , and that $\sum_{n=1}^{\nu} N_n(B) < \infty$, $B \in \mathcal{B}$. Then $N' = \sum_{n=1}^{\nu} N_n$ is a point process on E' representing the superposition of the clusters. Suppose, in addition, that N_1, \dots, N_{ν} are position-dependent marks of X_1, \dots, X_{ν} with distribution

$$K(x, A) = P\{N_n \subset A \mid X_n = x, \nu \geq n\}, \quad \text{for } A \text{ Borel in } \mathcal{N}'.$$

where \mathcal{N}' is the set of all counting measures μ on E' . Then the Laplace functional of a single cluster is

$$\begin{aligned} L(f; x) &\equiv E\left\{\exp\left[-\int_E f(y) N_n(dy)\right] \mid X_n = x, v \geq n\right\} \\ &= \int_{\mathcal{N}'} \exp\left[-\int_E f(y) \mu(dy)\right] K(x, d\mu). \end{aligned}$$

Furthermore,

$$\begin{aligned} L_{N'}(f) &= E\left\{E\left[\exp\left(-\sum_{n=1}^v \int_E f(y) N_n(dy)\right) \mid N\right]\right\} \\ &= E\left\{\prod_{n=1}^v L(f; X_n)\right\} = E\left\{\exp\sum_{n=1}^v \log L(f; X_n)\right\}. \end{aligned}$$

Thus,

$$L_{N'}(f) = L_N(g) \quad \text{where } g(x) = -\log L(f; x).$$

A point process N' whose distribution has a Laplace functional of this form is called a cluster process of N whose clusters have distribution $K(x, d\mu)$. We call N' a Poisson cluster process when the parent process N is Poisson. In this case,

$$L_{N'}(f) = \exp\left\{-\int_E \int_{\mathcal{N}'} (1 - e^{-\int f(y) \mu(dy)}) K(x, d\mu) \Lambda(dx)\right\}$$

where Λ is the mean of N . Note that any point process of the form

$$N'(A) = \int_{ExN'} \mu(A) M(dx d\mu)$$

is a cluster process of N when M is a Marked point process on ExN' of N with location-dependent marks with distribution $K(x, d\mu)$. Note that a compound Poisson process is a special case of a Poisson cluster process.

Example 2.13. Neyman-Scott Cluster Process. Suppose N is a stationary Poisson process on $E = \mathbb{R}^d$ with rate λ whose point locations represent cluster centers. Associated with a center, say at x , there is a cluster of β points whose distances from x are independent and identically distributed and β is a

random variable independent of x and the distances. In other words, a single cluster process centered at x is equal in distribution to the mixed sample process $N^x = \sum_{k=1}^{\beta} \delta_{x+Y_k}$, where β, Y_1, Y_2, \dots are independent and each Y_k has the distribution F . Let N' denote the sum of these single clusters centered at the point locations of N . Then N' is a Poisson cluster process whose clusters have the Laplace functional (recall § 2.2)

$$\begin{aligned} L(f; x) &= E \left\{ \exp \left[- \int_E f(y) N^x(dy) \right] \right\} \\ &= G \left(\int_E e^{-f(x+y)} F(dy) \right) \end{aligned}$$

where $G(s) = E(s^{\beta})$.

See Cox and Isham (1980) for further discussion of this and other cluster processes. Also, Serfozo (1984a) shows that a Cox process may be a good approximation for a cluster process whose clusters and points within a cluster are sparse (this is analogous to the Poisson approximation for rare events as in Theorem 3.7).

3. CONVERGENCE IN DISTRIBUTION OF POINT PROCESSES

As in many areas of probability, the notion of convergence in distribution plays an important role in the theory and applications of point processes. The focus of this section is on the convergence in distribution of sums of independent point processes to Poisson and infinitely divisible processes and on related rates of convergence and Poisson approximations.

3.1 Basics of Convergence in Distribution

We begin with a short review. Suppose X, X_1, X_2, \dots are random variables that take values in E (think of these as general random elements not as locations of points). The sequence X_n converges in distribution to X , written $X_n \xrightarrow{d} X$, if $Ef(X_n) \rightarrow Ef(X)$ for every bounded continuous $f: E \rightarrow \mathbb{R}$ (i.e. the probability distribution of X_n converges weakly to that of X). A simple characterization is that $X_n \xrightarrow{d} X$ if and only if $P\{X_n \in A\} \rightarrow P\{X \in A\}$ for each $A \in \mathcal{E}$ with $P\{X \in \partial A\} = 0$ (here ∂A means the boundary of A).

The importance of convergence in distribution is manifest by the following basic results (see Billingsley (1968) for further properties of convergence in distribution).

Continuous Mapping Principle. If $X_n \xrightarrow{d} X$, then $f(X_n) \xrightarrow{d} f(X)$ for every continuous $f: E \rightarrow E'$.

Convergence of Expectations. Suppose $E = \mathbb{R}_+$ and $X_n \xrightarrow{d} X$. Then $EX_n \rightarrow EX < \infty$ if and only if the X_n are uniformly integrable: $\sup_n E[X_n I(|X_n| \geq x)] \rightarrow 0$ as $x \rightarrow \infty$.

We now return to point processes. For a point process N on E , we define $B_N = \{B \in \mathcal{B}: N(\partial B) = 0 \text{ a.s.}\}$ and define I_N as the set of all intervals in B_N . Let N, N_1, N_2, \dots be point processes on E . Here are some characterizations of the convergence $N_n \xrightarrow{d} N$, which is defined as above (see Theorem 4.2 in Kallenberg (1983)).

Theorem 3.1. The following statements are equivalent.

- (i) $N_n \xrightarrow{d} N$.
- (ii) $(N_n(I_1), \dots, N_n(I_k)) \xrightarrow{d} (N(I_1), \dots, N(I_k)), \quad I_1, \dots, I_k \text{ in } I_N$.
- (iii) $\int_E f(x) N_n(dx) \xrightarrow{d} \int_E f(x) N(dx), \quad f \in C_K$.
- (iv) $L_{N_n}(f) \rightarrow L_N(f), \quad f \in C_K$.

It is easy to see that when statement (ii) is true, then it also holds for any I_1, \dots, I_k in E with $N(\partial I_j) = 0$ a.s. Similarly, statement (iii) extends to bounded $f: E \rightarrow \mathbb{R}_+$ with $\{x: f(x) > 0\} \in \mathcal{B}$ and $N(D_f) = 0$ a.s., where D_f is the set of discontinuity points of f . Condition (iv) is commonly used to establish (i), while (ii) and (iii) are used as properties or consequences of $N_n \xrightarrow{d} N$.

An elementary type of convergence is if N_n is a sequence of point processes defined by parameters α_n (e.g. a vector of numbers and measures), and N_n converges if α_n does. For example, if N_n is a Cox process directed by Λ_n and $\Lambda_n \xrightarrow{d} \Lambda$, then $N_n \xrightarrow{d} N$ where N is a Cox process directed by Λ . Similar statements apply to Poisson, negative binomial and sample processes. Also, when $E = \mathbb{R}_+$, then from Remark 1.2 it follows that $N_n \xrightarrow{d} N$ if and only if their point locations or interpoint locations converge accordingly

$((T_{n1}, \dots, T_{nk}) \xrightarrow{d} (T_1, \dots, T_k) \text{ for each } k)$. In particular, renewal processes

$N_n \xrightarrow{d} N$ if and only if their waiting time distributions converge accordingly ($F_n(x) \rightarrow F(x)$ for each F -continuity point x).

The following is a useful characterization of convergence when the limit process is simple.

Theorem 3.2. Suppose N is simple and

$$(3.1) \quad \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} P\{N_n(B) > m\} = 0, \quad B \in \mathcal{B}.$$

Then $N_n \xrightarrow{d} N$ if and only if

$$\lim_{n \rightarrow \infty} P\{N_n(B) = 0\} = P\{N(B) = 0\}, \quad B \in \mathcal{B}_N.$$

Proof. The necessity follows from Theorem 3.1, and the sufficiency follows from the uniqueness property in Remark 1.2(b) (since (3.1) implies that N_n is relatively compact; see §4 in Kallenberg (1983)).

Remark 3.3. A sufficient condition for (3.1) is $\limsup_{n \rightarrow \infty} EN_n(I) \leq EN(I) < \infty$,

$I \in \mathcal{I}_N$ (see §4 in Kallenberg (1983)).

3.2. Convergence to Poisson Processes

We now show that Poisson processes arise as limits of sums of sparse point processes and as limits of rare-event processes.

Suppose, for each n , that N_{n1}, N_{n2}, \dots is a finite or infinite sequence of independent point processes on E such that $\sum_j N_{nj}(B) < \infty$, $B \in \mathcal{B}$, a.s. and that they are uniformly null:

$$\lim_{n \rightarrow \infty} \sup_j P(N_{nj}(I) \geq 1) = 0, \quad I \in \mathcal{I}.$$

We call $\{N_{nj}\}$ a null array of point processes.

The following result is due to Crigelionis (1963) (he proved it for $E=\mathbb{R}$ and Jagers (1972) proved it for general spaces; see pp 175-6 in Kallenberg (1983) for further credits).

Theorem 3.4. Suppose $\{N_{nj}\}$ is a null array and N is a Poisson process with mean Λ . Then $\sum_j N_{nj} \xrightarrow{d} N$ if and only if

$$(3.2) \quad \lim_{n \rightarrow \infty} \sum_j P\{N_{nj}(B) \geq 1\} = \Lambda(B), \quad B \in \mathcal{B}_N, \text{ and}$$

$$(3.3) \quad \lim_{n \rightarrow \infty} \sum_j P\{N_{nj}(B) \geq 2\} = 0 \quad B \in \mathcal{B}.$$

Proof. This is a special case of Theorem 3.8. To illustrate the analysis dealing with convergence, we shall apply Theorem 3.2 to prove that (3.2) and (3.3) imply $N_n \xrightarrow{d} N$ when N is simple. Fix $B \in \mathcal{B}$ and let $p_{nj} = P\{N_{nj}(B) \geq 1\}$ and $N_n = \sum_j N_{nj}$. We can write,

$$(3.4) \quad P\{N_n(B) = 0\} = \prod_j P\{N_{nj}(B) = 0\} = \exp \sum_j \log P\{N_{nj}(B) = 0\}$$

$$= \exp\left(-\sum_j p_{nj} - \sum_j \sum_{k=1}^{\infty} p_{nj}^k/k\right).$$

Now for n large enough so that $\sup_j p_{nj} \leq 1/2$, we have

$$\sum_j \sum_{k=1}^{\infty} p_{nj}^k/k \leq \sum_j p_{nj}^2 \sum_{k=1}^{\infty} p_{nj}^k \leq 2 \sup_j p_{nj} \sum_j p_{nj} \rightarrow 0.$$

Using this and the assumption $\sum_j p_{nj} \rightarrow \Lambda(B)$ in (3.4) yields

$$P\{N_n(B) = 0\} \rightarrow e^{-\Lambda(B)} = P\{N(B) = 0\}, \quad B \in \mathcal{B}_N.$$

Next, observe that

$$\begin{aligned}
P\{N_n(B) \geq k\} &\leq P\left\{\bigcup_j \{N_{nj}(B) \geq 2\} \cup \{\text{at least } k \text{ different } N_{nj}(B) = 1\}\right\} \\
&\leq \sum_j P\{N_{nj}(B) \geq 2\} + \sum_{j_1 < \dots < j_k} p_{nj_1} \dots p_{nj_k} \\
&\leq \sum_j P\{N_{nj}(B) \geq 2\} + \left(\sum_j p_{nj}\right)^k \rightarrow \Lambda(B)^k.
\end{aligned}$$

Since $\Lambda(B) < \infty$, we can choose a finite partition $B = \bigcup_{i=1}^m B_i$ with $\Lambda(B_i) < 1$ for each i . Then using $P\{N_n(B) \geq k\} \leq \sum_{i=1}^m P\{N_n(B_i) \geq k\}$ and the preceding limit, we have

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} P\{N_n(B) \geq k\} \leq \lim_{k \rightarrow \infty} \sum_{i=1}^m \Lambda(B_i)^k = 0.$$

Thus Theorem 3.2 yields $N_n \xrightarrow{d} N$.

The preceding result is frequently used to justify a Poisson model for sums of sparse processes. Care should be taken in invoking this result, however, since the sum $\sum_j N_{nj}$ may converge to limits that are not Poisson; see Theorem 3.8. The closeness of $\sum_j N_{nj}$ to being Poisson can sometimes be assessed by results as in § 3.4. A corollary of the preceding for sums of renewal processes is as follows.

Corollary 3.5. Suppose N_{n1}, N_{n2}, \dots are independent renewal processes with interpoint distributions F_{n1}, F_{n2}, \dots , respectively, such that

$$(3.5) \quad \lim_{n \rightarrow \infty} \sup_j F_{nj}(t) = 0, \quad t \geq 0.$$

Let N be a Poisson process on R_+ with mean Λ . Then $\sum_j N_{nj} \xrightarrow{d} N$ if and only if $\sum_j F_{nj}(t) \rightarrow \Lambda(0, t]$, for each t with $\Lambda(\{t\}) = 0$.

Proof. By Theorem 3.4, it suffices to show that (3.5) is equivalent to (3.2) and (3.3). But this follows immediately from the relations

$$\sum_j P\{N_{nj}(0, t] \geq 1\} = \sum_j F_{nj}(t)$$

$$\sum_j P\{N_{nj}(0, t] \geq 2\} = \sum_j F_{nj} * F_{nj}(t) \leq \sum_j F_{nj}(t)^2 \leq \sup_t F_{nj}(t) \sum_j F_{nj}(t).$$

Example 3.6. Suppose N_1, N_2, \dots are independent renewal processes with interpoint distribution F . Consider the sum $\sum_{j=1}^n N_j$. As $n \rightarrow \infty$, this sum tends to infinity. However, we would like to normalize this sum, as in a central limit phenomenon, so that it converges to some point process. A natural way to do this is by rescaling the time axis. Accordingly, consider the process

$$N_n(0, t] \equiv \sum_{j=1}^n N_j(0, t/n]$$

which is the sum with the time axis expanded such that $1/n$ is the new unit of time. Suppose the derivative $F'(0) = \lambda > 0$. Then clearly

$N_{nj}(0, t] \equiv N_j(0, t/n]$ is a renewal process with interpoint distribution

$F(t/n)$, and the conditions of Corollary 3.5 are satisfied with $\sum_{j=1}^n$

$P\{N_j(0, t/n] \geq 1\} = n F(t/n) \rightarrow t\lambda$ as $n \rightarrow \infty$. Thus $N_n \xrightarrow{d} N$, a stationary Poisson process with rate λ .

The Poisson process is sometimes referred to as a point process of rare events because of the classical result that the number of successes in a sequence of Bernoulli trials is asymptotically Poisson as the probability of success tends to 0. We shall now present this rareness property for point processes other than Bernoulli processes. We discuss rareness in terms of thinning (recall §2.3).

Let N be a point process on $E = \mathbb{R}^d$. Suppose N'_n is a p_n -thinning of N , where $p_n \rightarrow 0$ (p_n is the probability of retaining a point). As $p_n \rightarrow 0$, the

points of N_n are rarer and $N'_n \xrightarrow{d} 0$. We would like to normalize N'_n to study its asymptotic behavior. Accordingly, consider the process $N_n(B) = N'_n(p_n^{-1}B)$, which is the thinned process with the space R^d compressed such that p_n^{-1} is the new unit of volume. In addition, suppose that N satisfies the weak law of large numbers

$$N(B_k)/|B_k| \rightarrow \lambda \quad \text{as } k \rightarrow \infty$$

for any sequence B_k in B with the area $|B_k| \rightarrow \infty$, where λ is a positive constant.

Theorem 3.7. Under the preceding assumptions, $N_n \xrightarrow{d} N^*$, where N^* is a stationary Poisson process with rate λ .

Proof. One approach is to verify $L_{N_n}(f) \rightarrow L_{N^*}(f)$ by direct computations. An alternative, informative approach is to observe that this assertion is a "randomized version" of Theorem 3.4. Namely, suppose $N = \sum_{j=1}^v \delta_{X_j}$, and let Z_{n1}, Z_{n2}, \dots be independent Bernoulli random variables with $p_n = P\{Z_{nj} = 1\}$. Then we can write

$$N_n(B) = N'(p_n^{-1}B) = \sum_{j=1}^v Z_{nj} 1(X_j \in p_n^{-1}B) = \sum_{j=1}^v N_{nj}(B),$$

where $N_{nj} = Z_{nj} \delta_{p_n X_j}$. Now, note that, given N , the $\{N_{nj}\}$ satisfy the assumptions of Theorem 3.4 (in terms of probabilities conditioned on N) with

$$\sum_{j=1}^v P\{N_{nj}(B) \geq 1 \mid N\} = p_n N(p_n^{-1}B) \xrightarrow{d} \lambda |B|.$$

Consequently,

$$E\left\{\exp\left[-\int_E f(x)N_n(dx)\right] \mid N\right\} \xrightarrow{d} \exp\left[-\int_E (1 - e^{-f(x)})\lambda dx\right] = L_{N^*}(f).$$

Then taking expectations yields $L_{N_n}(f) \rightarrow L_{N^*}(f)$.

Rényi (1967) proved Theorem 3.7 for a renewal process N . A special case is where N is non-random with one point at each integer $1, 2, \dots$. Then N_n is a Bernoulli point process with probability p_n of a point at each of the locations $p_n, 2p_n, 3p_n, \dots$, where $p_n N(p_n^{-1}t) \rightarrow t$, and $N_n \xrightarrow{d} N^*$ is the classical Binomial-to-Poisson convergence. This result (which is a functional limit theorem) can be viewed as the Poisson analogue of Donsker's functional central limit theorem for sums with the Wiener process as a limit. The normalization $N_n'(p_n^{-1}B)$ is rather natural: a normalization of the form $a_n N_n(b_n B)$ is analogous to $(\sum_{k=1}^k Z_k - a_n)/b_n$ (note that b_n rescales the space in the former and rescales the quantity in the latter). Kallenberg (1983) extended Theorem 3.7 to general E and initial processes N that vary with n such that the limits are nonstationary Poisson or Cox processes. For further results on thinnings of point-processes, see Serfozo (1980, 1984a,b) and the references therein.

3.3 Convergence to Infinitely Divisible Point Processes

We saw that the sum $\sum_j N_{nj}$ of a null array of point processes may converge in distribution to a Poisson process. Other limits, however, are possible. We now show that such sums may converge to infinitely divisible point processes and that these are the only possible limits.

A point process N on E is infinitely divisible if for each n there are independent identically distributed point processes N_1, \dots, N_n such that $N \stackrel{d}{=} N_1 + \dots + N_n$. We shall also be interested in point processes that are infinitely divisible with independent increments. The following are useful

characterizations for such processes (these are from Theorems 7.1, 7.2 and 6.1, respectively, in Kallenberg (1983)).

(a) A process N has independent increments if and only if it has a

representation of the form $N = N' + \sum_{j=1}^k Z_j \delta_{x_j}$, where N' is infinitely

divisible with independent increments, x_j are non-random points in E , $k \leq \infty$ is non-random, and Z_j are positive integer-valued random variables independent of N' .

(b) A point process N is infinitely divisible with independent increments if and only if

$$(3.6) \quad L_N(f) = \exp \left\{ - \sum_{m=0}^{\infty} \int_E (1 - e^{-mf(x)}) \gamma_m(dx) \right\}$$

where γ_m is a measure on E . For example, when $\gamma_m(dx) = F(\{m\})\Lambda(dx)$, then N is a compound Poisson process-whose Poisson points have mean Λ and mass distribution F .

(c) A point process N is infinitely divisible if and only if

$$(3.7) \quad L_N(f) = \exp \left\{ - \int_{\mathcal{N}_0} \left[1 - e^{-\int f(x)\mu(dx)} \right] \eta(d\mu) \right\}$$

where η is a measure on the set \mathcal{N}_0 of non-zero counting measures μ on E such that $\int_{\mathcal{N}_0} (1 - e^{-\mu(B)}) \eta(d\mu) < \infty$, $B \in \mathcal{B}$. In other words, N is infinitely divisible

if and only if it is a Poisson cluster process (recall § 2.5).

The following are fundamental results on the convergence of sums; see Theorems 6.1 and 7.2 of Kallenberg (1983). Suppose that $\{N_{nj}\}$ is a null array of point processes on E . Let N' be an infinitely divisible point process on E with independent increments as in (b) and let N^* be an infinitely divisible point process on E as in (c).

Theorem 3.8. If $\sum_j N_{nj} \xrightarrow{d} N$ to some N , then N is infinitely divisible. A necessary and sufficient condition for $\sum_j N_{nj} \xrightarrow{d} N^*$ is that

$$\lim_{n \rightarrow \infty} \sum_j P\{N_{nj}(I) = m\} = \gamma_m(I), \quad I \in I_N.$$

A necessary and sufficient condition for $\sum_j N_{nj} \xrightarrow{d} N^*$ is that

$$\lim_{n \rightarrow \infty} \sum_j \left[1 - L_{N_{nj}}(f) \right] = \int_{\mathcal{N}_0} \left[1 - e^{-\int f(x) \mu(dx)} \right] \eta(d\mu).$$

This condition is equivalent to

$$\lim_{n \rightarrow \infty} \sum_j P\{N_{nj}(I_1) = m_1, \dots, N_{nj}(I_r) = m_r\} = \eta\{\mu \in \mathcal{N}_0 : \mu(I_1) = m_1, \dots, \mu(I_r) = m_r\}$$

for I_1, \dots, I_r in I_N .

The preceding results show that infinitely divisible point processes arise naturally as Poisson cluster processes and as sums of sparse point processes. They also arise as limits of compound point processes with uniformly small masses, similar to the Poisson rare-event property in Theorem 3.7; see §8 of Kallenberg (1983) and Serfozo (1984). The convergence of point processes is an important tool for analyzing high-level exceedances or rare events of a stochastic process. Here is an elementary example.

Example 3.9. High-Level Exceedances of a Random Walk. Let Y_1, Y_2, \dots be a simple random walk on $\{0, 1, \dots\}$ whose probabilities of moving forward or backward one unit are respectively p and $q=1-p$, and $P\{Y_{k+1} = 1 \mid Y_k = 0\} = 1$. Consider the point process

$$N'_n(t) = \sum_{k=0}^{\infty} 1(Y_k \geq n) \delta_k[0, t], \quad t \geq 0.$$

This records the number of steps at which the walk exceeds the level n (or the amount of time the walk spends in $[n, \infty)$). We are interested in the behavior of N'_n as the level $n \rightarrow \infty$. Accordingly, we consider the normalized process $N_n(t) = N'_n(a_n t)$, where $a_n = (q/p)^n / (q/p - 1)$. As an application of Theorem 3.8, it follows that $N_n \xrightarrow{d} N$, where N is a compound Poisson process with stationary Poisson points at rate $(1-p/q)/2$ and mass distribution

$$F(\{m\}) = \sum_{k=1}^{\infty} g^{k*}(m) (p/q)^{m-1} (1-p/q)$$

where

$$g(2k-1) = (-1)^{k-1} \binom{1/2}{k} (4pq)^k / (2p) \quad k \geq 1.$$

The limit N is compound Poisson rather than just Poisson because there is a clumping of points in N_n : when the walk exceeds n it typically stays there for several steps in succession before wandering below n . This result appears in Serfozo (1980). See this paper, Leadbetter et. al (1983), Hsing et. al. (1987) and their references for similar results on high-level exceedances, or level-crossings for stationary and other processes.

3.4 Poisson Approximations

Regarding the preceding convergence theorems, a typical concern is the rate of the convergence or the nearness of the distribution of $\sum_{j=1}^n N_{nj}(A)$ to that of its limit. There are several metrics or distance measures available

for addressing such issues. We shall discuss the total-variation metric and its use for assessing the quality of Poisson approximations. The following results are based on ideas in Serfling (1975), Serfozo (1986) and their references.

We begin with an introduction to total-variation distances. Suppose X and Y are real-valued random variables not necessarily on the same probability space. The total variation distance between X and Y (or between their distributions) is

$$d(X,Y) \equiv \sup_B |P\{X \in B\} - P\{Y \in B\}|,$$

where the supremum is over all Borel sets B in \mathbb{R} . When X and Y are integer-valued, this reduces to

$$d(X,Y) = \frac{1}{2} \sum_m |P\{X=m\} - P\{Y=m\}|.$$

Here are some basic properties of this distance measure.

(a) Coupling Inequality: $d(X,Y) \leq P\{X' \neq Y'\}$, for any random variables X', Y' on a single probability space such that $X' \stackrel{d}{=} X$ and $Y' \stackrel{d}{=} Y$. One can derive bounds for $d(X,Y)$ by constructing X', Y' for which $P\{X' \neq Y'\}$ is small and calculable. There are X', Y' for which equality obtains.

(b) Triangle Inequality: $d(X,Y) \leq d(X,Z) + d(Z,Y)$.

(c) $d(X,Y) \leq E[\sup_B |P\{X' \in B|U\} - P\{Y' \in B|V\}|]$, for any random variables

X', Y', U, V on a single probability space such that $X' \stackrel{d}{=} X$ and $Y' \stackrel{d}{=} Y$. We write this as $d(X,Y) \leq E[d(X'|U, Y'|V)]$.

(d) If X_1, \dots, X_n are independent, and Y_1, \dots, Y_n are independent, then

$$d\left(\sum_{j=1}^n X_j, \sum_{j=1}^n Y_j\right) \leq \sum_{j=1}^n d(X_j, Y_j).$$

(e) If X and Y are Poisson with respective means α and β , then $d(X, Y) \leq |\alpha - \beta|$.

(f) If X is a Bernoulli random variable with $p = P\{X=1\}$, and Y is Poisson with mean p , then $d(X, Y) \leq p^2$.

(g) Measurable Mapping Principle: If $d(X_n, X) \rightarrow 0$, then $X_n \xrightarrow{d} X$ and, moreover, $f(X_n) \xrightarrow{d} f(X)$ for any (measurable) $f: \mathbb{R} \rightarrow \mathbb{R}$ (cf. the continuous mapping principle in §3.1). This observation is useful for establishing limit theorems.

The next result, as we shall show, is useful for assessing Poisson approximations. Similar results hold for sums of dependent variables and for compound Poisson approximations.

Theorem 3.10. Suppose X_1, \dots, X_n are independent non-negative integer-valued random variables and Y is a Poisson random variable with mean α . Then

$$d\left(\sum_{j=1}^n X_j, Y\right) \leq \sum_{j=1}^n [P\{X_j \geq 2\} + P\{X_j \geq 1\}^2] + \left|\alpha - \sum_{j=1}^n P\{X_j \geq 1\}\right|.$$

Proof: Let Y_1, \dots, Y_n be independent Poisson random variables with respective means $P\{X_1 \geq 1\}, \dots, P\{X_n \geq 1\}$. Then by properties (b), (d), (e), we have

$$\begin{aligned}
d\left(\sum_{j=1}^n X_j, Y\right) &\leq d\left(\sum_{j=1}^n X_j, \sum_{j=1}^n 1(X_j \geq 1)\right) + d\left(\sum_{j=1}^n 1(X_j \geq 1), \sum_{j=1}^n Y_j\right) \\
&\quad + d\left(\sum_{j=1}^n Y_j, Y\right) \\
&\leq \sum_{j=1}^n d(X_j, 1(X_j \geq 1)) + \sum_{j=1}^n d(1(X_j \geq 1), Y_j) \\
&\quad + \left|\alpha - \sum_{j=1}^n P\{X_j \geq 1\}\right|.
\end{aligned}$$

Using $d(X_j, 1(X_j \geq 1)) = P\{X_j \geq 2\}$ and $d(1(X_j \geq 1), Y_j) \leq P\{X_j \geq 1\}^2$ from (f) in the last line yields the assertion.

We now apply the preceding ideas to point processes. Here is a companion to the Poisson convergence in Theorem 3.4.

Corollary 3.11. Suppose $\{N_{nj}\}$ is a null array of point processes on E and N is a Poisson process on E with mean Λ . Then, for $A \in \mathcal{E}$,

$$\begin{aligned}
d\left(\sum_j N_{nj}(A), N(A)\right) &\leq \sum_j P\{N_{nj}(A) \geq 2\} + \sum_j P\{N_{nj}(A) \geq 1\}^2 \\
&\quad + \left|\Lambda(A) - \sum_j P\{N_{nj}(A) \geq 1\}\right|.
\end{aligned}$$

This is an immediate consequence of Theorem 3.10 (which also holds for $n=\infty$). Under the assumptions of Theorem 3.10, the right side of the preceding inequality tends to 0 as $n \rightarrow \infty$. One is typically interested in the case when N has the mean $\Lambda(A) = \sum_j P\{N_{nj}(A) \geq 1\}$; then the right side consists of only the first two terms.

The next result is a companion to the rare-event property of Theorem 3.7.

Corollary 3.12. Suppose N_n is the rescaled p_n -thinning of N as in Theorem 3.7 and N^* is a stationary Poisson process with rate λ . Then, for $B \in \mathcal{B}$,

$$(3.8) \quad d(N_n(B), N^*(B)) \leq p_n^2 E[N(p_n B)] + E[|\lambda|B| - p_n N(p_n B)|].$$

Proof. Using property (c), Theorem 3.10 and the notation in the proof of Theorem 3.7, we have

$$\begin{aligned} d(N_n(B), N^*(B)) &\leq E[d(N_n(B) | N, N^*(B))] \leq E\left[\sum_{j=1}^v P\{Z_{nj} 1(p_n X_n \in B) = 1 | N\}^2\right] \\ &\quad + E\left[|\lambda|B| - \sum_{j=1}^v P\{Z_{nj} 1(p_n X_n \in B) = 1\}|\right] \\ &= E\left[\sum_{j=1}^v p_n^2 1(p_n X_n \in B)\right] + E\left[|\lambda|B| - \sum_{j=1}^v p_n 1(p_n X_n \in B)\right], \end{aligned}$$

where the last two expectations are equal to those in the assertion.

Note that the right side of (3.8) tends to zero when N satisfies the weak law of large numbers as in Theorem 3.7 and $\limsup_{n \rightarrow \infty} p_n E[N(p_n B)] < \infty$ (consequently, $N_n(B) \xrightarrow{d} N^*(B)$, which we also know from Theorem 3.7).

4. RENEWAL THEORY

The theory of renewal processes was developed in the '40s, '50s and '60s by Blackwell, Doob, Feller, Smith and others. The major topics of this subject are: (a) The key renewal theorem, which describes the limiting behavior of the solution to a renewal equation. (b) Applications of the key renewal theorem to obtain limits of means and distributions of functionals of renewal and regenerative processes. (c) Limit laws of renewal, compound renewal or regenerative processes that are consequences of analogous limit laws for sums of independent random variables. (d) Processes with a renewal-like structure (e.g. alternating, transient, or branching renewal processes). (e) Statistical properties of renewal and regenerative processes. (f) Applications in systems that regenerate over time (e.g. systems involving queueing, reliability, inventory control or cash flows).

Since renewal theory is a common topic in introductory texts on stochastic processes, our coverage will be brief and confined to only the first three topics, with a novel treatment of the last two. Basic references are Feller (1971), Çinlar (1975) and Gut and Prabhu (1987).

4.1 Distributions of Renewal Processes

Suppose that N is a renewal process on R_+ with renewal times $0 < T_1 < T_2 < \dots$ and independent waiting times between renewals W_1, W_2, \dots that have the distribution F . For simplicity, we write $N_t = N[0, t]$. Some useful relations between the numbers of renewals and the renewal times are

$$N_t = \sum_{n=1}^{\infty} 1(T_n \leq t) = \sum_{n=1}^{\infty} n 1(T_n \leq t < T_{n+1}) = \sup\{n: T_n \leq t\},$$

$$\{N_t \geq n\} = \{T_n \leq t\}, \text{ and } T_{N_t} \leq t < T_{N_t+1}.$$

These properties, which follow by inspection, are true for any point process N on R_+ with $N(R_+) = \infty$. An immediate consequence is

$$\begin{aligned}
 P\{N_t = n\} &= P\{N_t \geq n\} - P\{N_t \geq n+1\} \\
 &= P\{T_n \leq t\} - P\{T_{n+1} \leq t\} = F^{n*}(t) - F^{(n+1)*}(t),
 \end{aligned}$$

where F^{n*} is the n -th fold convolution of F . In addition, all moments of N_t are finite and

$$EN_t = \sum_{n=1}^{\infty} E[1(T_n \leq t)] = \sum_{n=1}^{\infty} F^{n*}(t).$$

The renewal function

$$U(t) \equiv \sum_{n=0}^{\infty} F^{n*}(t) = 1 + EN_t,$$

plays an important role in characterizing N and it is the focus of the key renewal theorem. We first note that U and the distribution of N uniquely determine each other. To see this, take the Laplace transform of the preceding equation to get $\tilde{U}(s) = (1 - \tilde{F}(s))^{-1}$, where \tilde{U} , \tilde{F} are Laplace transforms of U, F . Then clearly \tilde{U} and \tilde{F} uniquely determine each other and hence so do U and F . But we know by Remark 1.11 that F and the distribution of N uniquely determine each other. Thus the assertion follows. As an example, N is a stationary Poisson process with rate μ^{-1} if and only if $U(t) = 1 + t/\mu$. Unfortunately, nice expressions for renewal functions are the exception rather than the rule.

The following limit laws for N_t describe its ascension to ∞ as $t \rightarrow \infty$. We let $\mu \leq \infty$ denote the mean of F , and we interpret $1/\infty$ as 0. Recall that a distribution F is arithmetic if its jumps are concentrated on points of the form $d, 2d, 3d, \dots$ and the largest d with this property is the span.

Strong Law of Large Numbers: $N_t/t \rightarrow 1/\mu$ a.s. (Corollary 4.12).

Central Limit Theorem: If the variance σ^2 of F is finite, then

$(N_t - t/\mu)/(\sigma t^{1/2} \mu^{-3/2}) \xrightarrow{d} Z$, where Z is a standard normal random variable.

Convergence of Moments: $E(N_t/t)^r \rightarrow (1/\mu)^r$ for $r \geq 1$, and, when σ^2 exists,

$$EN_t = t/\mu + (\sigma^2 - \mu^2)/(2\mu^2) + o(1) \quad \text{and} \quad \text{Var}N_t = t\sigma^2/\mu^3 + o(t),$$

where $t \rightarrow \infty$ and, in case F is arithmetic, then t is a multiple of the span (these statements follow from Example 4.8).

Renewal processes are intimately related to processes that are regenerative or have regenerative increments as follows. Suppose $X = \{X_t; t \in R_+\}$ is a stochastic process with state space E and sample paths in the set $D(R_+, E)$ of all functions from R_+ to E that are right-continuous and have left-hand limits. The X is a regenerative process with regeneration times T_1, T_2, \dots if $N = \sum_{n=1}^{\infty} \delta_{T_n}$ is a renewal process and, for each t_1, \dots, t_k , w in R_+ and A_1, \dots, A_k in E ,

$$(4.1) \quad P\{W_{n+1} \leq w, X_{T_n+t_1} \in A_1, \dots, X_{T_n+t_k} \in A_k \mid W_1, \dots, W_n; X_s, s \leq T_n\} \\ = P\{W_1 \leq w, X_{t_1} \in A_1, \dots, X_{t_k} \in A_k\}, \quad n = 0, 1, \dots$$

The n th cycle of X consists of the information

$$\xi_n = (W_n, \{X_{T_{n-1}+u}; 0 \leq u \leq W_n\}), \quad n = 1, 2, \dots$$

which is the cycle time $W_n = T_n - T_{n-1}$ and the trajectory of X in $[T_{n-1}, T_n]$. Condition (4.1) is equivalent to ξ_1, ξ_2, \dots being independent and identically distributed (which also implies that N is a renewal process). This definition is equivalent to that in Çinlar (1975); we avoid his assumption that T_n is a stopping time of X by including W_{n+1} in (4.1). Positive recurrent Markov or semi-Markov processes are examples of regenerative processes.

Now suppose $X = \{X_t; t \in R_+\}$ has state space R and sample paths in $D(R_+, R)$. The X has regenerative increments over T_1, T_2, \dots if $N = \sum_{n=1}^{\infty} \delta_{T_n}$ is a renewal process and, for each w, x_1, \dots, x_k and $0 \leq s_1 < t_1 < \dots < s_k < t_k$,

$$(4.2) \quad P\{W_{n+1} \leq w, X_{T_n+t_1} - X_{T_n+s_1} \leq x_1, \dots, X_{T_n+t_k} - X_{T_n+s_k} \leq x_k \\ \mid W_1, \dots, W_n; X_s, s \leq T_n\}$$

$$= P\{W_1 \leq w, X_{t_1} - X_{s_1} \leq x_1, \dots, X_{t_k} - X_{s_k} \leq x_k\}, \quad n = 0, 1, \dots$$

This condition has an interpretation similar to (4.1) via cycle information. Such processes are useful for modeling additive functionals of regenerative processes as in §4.3 and §4.4.

The notion of regeneration is also manifest in point processes. A point process N on R_+ is regenerative over $T_1 < T_2 < \dots$ if $\{N_t; t \in R_+\}$ has regenerative increments over these times. The T_n 's need not necessarily be points of N . For example, the times at which a regenerative process enters a certain set form a regenerative point process. Also, a point process N on R_+ has regenerative interpoint distances if W_1, W_2, \dots is a regenerative sequence over discrete times $v_1 < v_2 < \dots$. See Neuts (1979) and Wold (1948) for related point processes (the N is a Wold process when W_1, W_2, \dots is Markovian). The analysis of such point processes is similar in part to that of renewal processes.

4.2 Key Renewal Theorem

We now consider the renewal equation

$$f(t) = g(t) + \int_{[0,t]} f(t-s)F(ds), \quad t \geq 0,$$

where F is the distribution as above and f and g are functions from R_+ to R_+ that are bounded on finite intervals. The F and g are known and f is unknown. The f is typically the mean or distribution of a functional of N as we discuss below. The assumption that f, g are non-negative is for convenience; one can treat real-valued f, g by considering their positive and negative parts separately (the positive and negative parts of f are $f^+(x) = \max\{0, f(x)\}$ and $f^-(x) = -\min\{0, f(x)\}$). The renewal equation $f = g + F * f$ has the unique solution

$$f(t) = \int_{[0,t]} g(t-s)U(ds) = U * g(t).$$

This is clearly a solution since $g + (U * g) * F = U * g$. See Chapter XI of Feller (1971) for a proof of the uniqueness and for proofs of what follows.

A major issue of renewal theory is the existence of the limit of $U * g(t)$ as $t \rightarrow \infty$. To describe this, we need one more concept. A function $g: R_+ \rightarrow R_+$ is directly Riemann integrable (denoted $g \in D$) if the following sums exist

$$s(\delta) = \delta \sum_{n=1}^{\infty} \inf \{g(x) : (n-1)\delta \leq x \leq n\delta\}$$

$$S(\delta) = \delta \sum_{n=1}^{\infty} \sup \{g(x) : (n-1)\delta \leq x \leq n\delta\}$$

and $\lim_{\delta \rightarrow 0} (S(\delta) - s(\delta)) = 0$. Consequently, g is bounded on finite intervals

and

$$\lim_{\delta \rightarrow 0} s(\delta) = \lim_{\delta \rightarrow 0} S(\delta) = \int_0^{\infty} g(x)dx,$$

which is the usual indefinite Riemann integral (the preceding conditions are stronger than those needed for the existence of the integral). Clearly $g \in D$ if the number of discontinuities of g is finite in each finite interval and either (i) g is zero outside a finite interval or (ii) g is decreasing and $\int_0^{\infty} g(x)dx < \infty$. Here is another characterization.

Lemma 4.1. (Hinderer (1986)). $g \in D$ if and only if $S(\delta) < \infty$ for some δ and g is continuous almost everywhere with respect to Lebesgue measure.

The following major results describe the limiting behavior of $U * g$ and U . (Recall that $1/\infty = 0$.)

Key Renewal Theorem 4.2. If F is not arithmetic and $g \in D$, then

$$\lim_{t \rightarrow \infty} U * g(t) = \mu^{-1} \int_0^{\infty} g(x)dx.$$

If F is arithmetic with span d and the following sum is finite, then

$$\lim_{n \rightarrow \infty} U * g(x+nd) = \mu^{-1} \sum_{k=1}^{\infty} g(x+kd).$$

Blackwell's Renewal Theorem 4.3. If F is not arithmetic, then $U(t+h) - U(t) \rightarrow h/\mu$ as $t \rightarrow \infty$, $h > 0$. If F is arithmetic, then this limit holds with t a multiple of the span.

These two limit theorems are actually equivalent: The latter follows from the former with $g(t) = 1(t \in (0, u])$, and the proof of the reverse implication is implicit in Feller's (1971) proof of the key renewal theorem. See Çinlar (1972), Kesten (1974) and their references for generalizations of the key renewal theorem to Markov renewal processes.

4.3 Applications of the Key Renewal Theorem

The key renewal theorem yields limit theorems for expectations of functionals of renewal and regenerative processes. We shall present a general approach for identifying new applications and review some standard examples.

We first note that the standard applications of the key renewal theorem, as in Feller (1971), are all limit statements of the form $\lim_{t \rightarrow \infty} E\phi(t, N, X)$.

Here $Z_t \equiv \phi(t, N, X)$ is a functional of t , the renewal process N and a random element X (e.g. a process or mark associated with N). The $\Gamma(t) \equiv EZ_t$ satisfies a renewal equation, and so $EZ_t = U * g(t)$ for some g . This raises the questions: What are the possible functionals ϕ for which EZ_t satisfies a renewal equation? Must Z be regenerative? What about X ? The following observation clears the air in this regard.

Observation 4.4. Suppose $Z = \{Z_t; t \in \mathbb{R}_+\}$ is a real-valued stochastic process on the same probability space as the renewal process N . Assume that

$f(t) \equiv EZ_t$ is finite for each t . Then f satisfies the renewal equation

$f = g + F * f$ with

$$(4.3) \quad g(t) \equiv E[Z_t 1(T_1 > t)] + \int_{[0, t]} [E(Z_t | T_1 = s) - EZ_{t-s}] F(ds).$$

Hence $EZ_t = U * g(t)$, $t \in \mathbb{R}_+$. Furthermore, when g^+ and g^- are in D and F is not arithmetic, then

$$\lim_{t \rightarrow \infty} EZ_t = \mu^{-1} \int_0^\infty g(x) dx.$$

An analogous limit holds when F is arithmetic.

Proof. This follows since

$$\begin{aligned} f(t) &= E[Z_t 1(T_1 > t)] + \int_{[0, t]} E(Z_t | T_1 = s) F(ds) - F * f(t) + F * f(t) \\ &= g(t) + F * f(t). \end{aligned}$$

This observation has the surprising implication that the mean of any real-valued process Z_t has the representation $EZ_t = U * g(t)$. This is of interest, of course, only when Z depends on N ; otherwise, it is vacuous. Hence, if the limit of EZ_t appears to exist, then it is a candidate for the key renewal theorem. The current literature on renewal theory suggests that one "set up" a renewal equation to obtain the representation $EZ_t = U * g(t)$ or that one derive it directly, for each application. This, however, is not needed: The preceding observation says that this representation is automatically satisfied for any application, and that g is given by (4.3).

Remark 4.5. Two special cases of (4.3) are:

$$(4.4) \quad g(t) = E[Z_t 1(T_1 > t)], \quad \text{when } Z \text{ is regenerative,}$$

$$(4.5) \quad g(t) = E(Z_{t \wedge T_1}), \quad \text{when } Z \text{ has regenerative increments.}$$

The following are examples of the key renewal theorem and the preceding comments. For simplicity, we assume hereafter that F is not arithmetic.

Example 4.6. Regenerative Processes. Suppose $X = \{X_t : t \in \mathbb{R}_+\}$ is a regenerative process over T_1, T_2, \dots and $h: E \rightarrow \mathbb{R}_+$. Then from Observation 4.4 and (4.4), we have

$$(4.6) \quad \lim_{t \rightarrow \infty} E h(X_t) = \mu^{-1} \int_0^{\infty} E[h(X_s) 1(T_1 > s)] ds,$$

provided the integral exists. It exists when $E[T_1 b(T_1)] < \infty$, where $b(t)$

$\equiv \sup_{s \leq t} E[h(X_s) \mid T_1 = t]$. The important case of (4.6) for $h(X_t) = 1(X_t \in A)$,

$A \in E$, reads

$$\lim_{t \rightarrow \infty} P\{X_t \in A\} = \mu^{-1} \int_0^{\infty} P\{X_s \in A, T_1 > s\} ds = \mu^{-1} E\left[\int_0^{T_1} 1(X_s \in A) ds\right].$$

Other special cases are in the next example.

Example 4.7. Backward and Forward Recurrence Times. Two basic functionals of the renewal process N are

$$Y_t \equiv t - T_{N_t} \quad \text{and} \quad Y'_t \equiv T_{N_t+1} - t,$$

the backward and forward recurrence times at t (the time since the last renewal before t and the time to the next one after t). Another functional

is $L_t \equiv W_{N_t} = Y'_t - Y_t$, the length of the renewal interval containing t . To

obtain limits of the means or joint distributions of these processes,

consider $Z_t = h(Y_t, Y'_t, L_t)$ where $h: R_+^3 \rightarrow R$. Assume that $E[T_1 b(T_1)] < \infty$, where $b(t) \equiv \sup_{s \leq t} |h(s, t-s, t)|$. The Z_t is regenerative, and so by Example 4.6, we

have

$$\lim_{t \rightarrow \infty} E h(Y_t, Y'_t, L_t) = \mu^{-1} E\left[\int_0^{T_1} h(s, T_1-s, T_1) ds\right].$$

Hinderer (1985) discusses this and related results. Here are some special cases:

$$P\{Y_t \leq y\} \rightarrow \mu^{-1} \int_0^y [1-F(s)] ds$$

(Y'_t and L_t have this same limiting distribution),

$$P\{Y_t > y, Y'_t > y'\} \rightarrow 1 - \mu^{-1} \int_0^{y+y'} [1-F(s)] ds$$

$$P\{Y_t/L_t \leq x\} \rightarrow x, \quad 0 \leq x \leq 1,$$

(the latter also holds with Y'_t in place of Y_t), and

$$E[Y_t^k Y_t^{\ell} L_t^m] \rightarrow E(T_1^{k+\ell+m+1}) / \left[\mu(k+\ell+1) \binom{k+\ell}{k} \right],$$

provided the last expectation is finite.

Example 4.8. Processes with Regenerative Increments. Suppose $X = \{X_t : t \in \mathbb{R}_+\}$ is an increasing process with regenerative increments over T_1, T_2, \dots . Assume that $a \equiv EX_{T_1}$ exists. By the strong law of large numbers for regenerative processes (Corollary 4.12(b)), we know that $t^{-1}X_t \rightarrow a/\mu$ a.s. How does $t^{-1}EX_t$ behave as $t \rightarrow \infty$? Assume that F has a variance σ^2 and that the distribution $G(t) \equiv a^{-1}E(X_{t \wedge T_1})$ has a finite mean α . Then

$$(4.7) \quad \lim_{t \rightarrow \infty} [EX_t - at/\mu] = (a/\mu^2)[(\mu^2 + \sigma^2)/2 - a\mu].$$

This follows by applying Observation 4.4 and (4.5) to the process

$Z_t \equiv X_t - at/\mu$, which has regenerative increments, and observing that the limit above equals

$$\mu^{-1} \int_0^\infty [E(X_{s \wedge T_1}) - (a/\mu)E(s \wedge T_1)] ds = (a/\mu^2) \left[\int_0^\infty [1 - E(s \wedge T_1)] ds - \int_0^\infty [1 - G(s)] ds \right],$$

which reduces to the right-hand side of (4.7). An immediate consequence of

(4.7) is $\lim_{t \rightarrow \infty} t^{-1}EX_t = a/\mu$. Special cases of this are the convergence of moments of N_t in §4.1 and Example 4.13.

4.4 Laws of Large Numbers

Strong laws of large numbers for renewal, Markov and regenerative processes appear frequently in operations research studies. Their main use is for obtaining easy-to-understand performance measures of systems. They are also useful for establishing objective functions or constraints in optimization problems. The literature contains a variety of laws of large numbers for point processes on \mathbb{R}_+ and for stochastic processes associated with them. We shall present a general limit law that yields many of these

ostensively different laws as corollaries. We discuss special uses for renewal, Markov and regenerative processes.

Suppose that $N = \sum_{n=1}^{\infty} \delta_{T_n}$ is a point process on R_+ and $Z = \{Z_t : t \in R_+\}$ is an increasing real-valued process associated with N . Since N_t is the generalized inverse of T_n (recall $N_{T_n} = n$), one would anticipate that the limiting behavior of N_t as $t \rightarrow \infty$ would be the inverse of that of T_n as $n \rightarrow \infty$. Also, the limiting behavior of Z_t should mimic that of the embedded process Z_{T_n} . The following results formalize these statements. Here we let a and μ denote positive constants, one of which may be infinite. Also, in the proofs we write the limit statements using the symbol \sim , where $g(x) \sim h(x)$ as $x \rightarrow \infty$ means $g(x)/h(x) \rightarrow 1$. We also suppress the a.s. For example, $n^{-1}T_n \rightarrow \mu$ a.s. becomes $T_n \sim n\mu$.

Theorem 4.9. If $n^{-1}T_n \rightarrow \mu$ a.s., then $t^{-1}Z_t \rightarrow a/\mu$ a.s. if and only if $n^{-1}Z_{T_n} \rightarrow a$ a.s.

Proof. If $Z_t \sim at/\mu$, then clearly $Z_{T_n} \sim aT_n/\mu \sim an$. Conversely, if $Z_{T_n} \sim an$, then since Z is increasing and $T_{N_t} \leq t < T_{N_t+1}$, we have

$$Z_t/t \leq Z_{T_{N_t+1}}/T_{N_t+1} \sim a(N_t+1)/(\mu N_t) \sim a/\mu,$$

and

$$Z_t/t \geq Z_{T_{N_t}}/T_{N_t} \sim aN_t/[\mu(N_t+1)] \sim a/\mu.$$

Combining these statements, yields $Z_t \sim at/\mu$.

Corollary 4.10. $t^{-1}N_t \rightarrow \mu^{-1}$ a.s. if and only if $n^{-1}T_n \rightarrow \mu$ a.s.

Proof. If $T_n \sim n\mu$, then by Theorem 4.9 with $Z_t = N_t$ and recognizing $n^{-1}N_{T_n} = 1$, we get $N_t \sim t/\mu$. Conversely, if $N_t \sim t/\mu$, then clearly $T_n = n(T_n/N_{T_n}) \sim n\mu$.

Remarks 4.11. (a) The preceding results are also true with a and μ random or with a.s. replaced by convergence in probability or in distribution. (b) Theorem 4.9 is useful for analyzing processes of the form $Z_t = Z_t^1 - Z_t^2$, where Z_t^1 and Z_t^2 are increasing (e.g. when Z_t has bounded variation). Just apply the result to each part separately. (c) Many versions of Little's law $L = \lambda W$ for queues are consequences of Theorem 4.9. (d) Corollary 4.10 yields a law of large numbers for any point process N whose interpoint distances satisfy such a law (i.e. when $n^{-1}(W_1 + \dots + W_n)$ converges).

We now specialize the results above to renewal and to regenerative processes.

Corollary 4.12. (a) If N is a renewal process with $\mu = ET_1 \leq \infty$, then $t^{-1}N_t \rightarrow \mu^{-1}$ a.s. (b) Suppose Z has regenerative increments over T_1, T_2, \dots and $Z_t = Z_t^1 - Z_t^2$, where Z^1 and Z^2 are increasing processes. If $a = EZ_{T_1}^1$ exists, and a and μ are not both infinite, then $t^{-1}Z_t \rightarrow a/\mu$ a.s.

Proof. By the strong law of large numbers for sums of independent variables, we know that $n^{-1}T_n \rightarrow \mu$ a.s. and $n^{-1}Z_{T_n}^1 \rightarrow a$ a.s. Thus (a) and (b) follow by Corollary 4.10 and Theorem 4.9, respectively.

Example 4.13. Additive Functionals of Regenerative and Markov Processes.

Suppose $X = \{X_t: t \in \mathbb{R}_+\}$ is a regenerative process over T_1, T_2, \dots . Consider the additive functional

$$(4.8) \quad Z_t = \int_0^t f(X_s) ds + \sum_{0 \leq s \leq t} g(X_{s-}, X_s),$$

where $f: E \rightarrow \mathbb{R}$ and $g: E^2 \rightarrow \mathbb{R}$. The $f(x)$ might be the cost per unit time of X

being in state x , and $g(x,y)$ might be the cost of the process jumping from x to y (assuming the sum is finite). Then Z_t would be the total cost up to time t . We can write $Z_t = Z_t^+ - Z_t^-$, where $Z_t^+(Z_t^-)$ are the positive and negative parts of Z_t (Z_t^\pm is given by (4.8) with f, g replaced by f^\pm, g^\pm).

Assume that $a \equiv EZ_{T_1} = EZ_{T_1}^+ - EZ_{T_1}^-$ exists. We can write

$$(4.9) \quad a = \int_0^\infty E[f(X_s)1(T_1 > s)]ds + \sum_{0 \leq s \leq t} E[g(X_{s-}, X_s)1(T_1 > s)] \\ = E\left[\int_0^{T_1} f(X_s)ds\right] + E\left[\sum_{s \leq T_1} g(X_{s-}, X_s)\right].$$

Then from Corollary 4.12(b) and Example 4.8, we have

$$(4.10) \quad t^{-1}Z_t \rightarrow a/\mu \text{ a.s. and } t^{-1}EZ_t \rightarrow a/\mu.$$

Expression (4.7) also applies.

In particular, suppose X is a Markov process with countable state space E and transition rates

$$q(x,y) \equiv \lim_{h \downarrow 0} P\{X(t+h) = y \mid X(t) = x\}/h, \quad x,y \in E.$$

Assume that X is irreducible and positive-recurrent with limiting distribution $\pi(x)$. Then (4.10) holds with (4.9) reduced to

$$a = \sum_x f(x)\pi(x) + \sum_x \pi(x) \sum_{y \neq x} q(x,y)g(x,y).$$

For instance, suppose $Z_t = \int_0^t 1(X_s = x)ds$, the amount of time X spends in state x up to time t . With no loss in generality, we assume that $X_0 = x$. Then $t^{-1}Z_t \rightarrow \pi(x)$ a.s. and $t^{-1}EZ_t \rightarrow \pi(x)$. Furthermore, from (4.7),

$$\lim_{t \rightarrow \infty} [EZ_t - t\pi(x)] = \pi(x)[(\mu^2 + \sigma^2)/(2\mu) - q(x)^{-1}],$$

where $q(x) = \sum_{y \neq x} q(x,y)$, so $q(x)^{-1}$ is the mean sojourn time of X in state x ,

and μ and σ^2 are the mean and variance of the time between two entrances into state x .

5. STATIONARY POINT PROCESSES

In this section we discuss a few basic properties of stationary point processes on \mathbb{R} . These concern their infinitesimal behavior and Palm distributions (a conditional distribution of a process given it has a point at a certain location). Our coverage does not include ergodic theorems, spectral analysis, and prediction and filtering. As for applications, stationary point processes with marks have been especially useful in characterizing non-Markovian stationary queueing processes (Franken et. al (1981) and Baccelli and Bremaud (1987)).

5.1. Definitions and Examples

A stochastic process $X = \{X_t : t \in \mathbb{R}\}$ with a general state space E is stationary if $(X_{t_1+h}, \dots, X_{t_n+h}) \stackrel{d}{=} (X_{t_1}, \dots, X_{t_n})$, for each h, t_1, \dots, t_n in \mathbb{R} .

That is, the distribution of X is invariant under translations of the time axis. A simple way of expressing this is that $\theta_h X \stackrel{d}{=} X$ for each $h \in \mathbb{R}$, where $\theta_h X$ is the process X with the time axis translated by h , i.e.

$\theta_h X(t) \equiv X(t+h)$, $t \in \mathbb{R}$. Stationarity of point processes is similar. A point process N on \mathbb{R} is stationary if

$(N(B_1 + h), \dots, N(B_n + h)) \stackrel{d}{=} (N(B_1), \dots, N(B_n))$ for each B_1, \dots, B_n in \mathcal{E} , $h \in \mathbb{R}$.

In other words, the increments of N are stationary or invariant in distribution under time translations. Simply stated, $\theta_h N \stackrel{d}{=} N$ for each $h \in \mathbb{R}$, where $\theta_h N(B) \equiv N(B + h)$, which is the process N with the time axis translated by h . This notion readily extends to point processes on \mathbb{R}_+ , \mathbb{R}^d or on other groups or semigroups. We shall restrict our discussion to processes on \mathbb{R} . Recall the convention that the point locations of N are labeled such that

$$\dots \leq T_{-1} \leq T_0 < 0 \leq T_1 \leq T_2 \leq \dots$$

We have already discussed some stationary point processes: Poisson and compound Poisson processes with constant rates, and Cox processes with stationary intensity processes. Here are some more examples.

Example 5.1. Stationary Renewal Processes. Let F be a distribution with $F(0) = 0$ and finite mean μ . Suppose N is a point process on \mathbb{R} with independent inter-point distances such that each W_n , $n \neq 1$, has the distribution F and $W_1 = T_1 - T_0$ is such that $-T_0$ and T_1 have the joint distribution

$$P\{-T_0 > u, T_1 > v\} = 1 - \mu^{-1} \int_0^{u+v} [1-F(t)]dt.$$

In particular,

$$P\{-T_0 \leq t\} = P\{T_1 \leq t\} = \mu^{-1} \int_0^t [1-F(s)]ds.$$

Then it follows that N is stationary. This N is called a stationary renewal process.

Example 5.2. Functions of Stationary Processes. Many stationary point processes arise as functions of stationary processes as follows. Suppose X is a stationary process with state space E and N is a point process defined by $N(A) = \phi(X, A)$ where $\phi: \mathcal{X} \times \mathcal{E} \rightarrow \{0, 1, \dots\}$ and \mathcal{X} is the set of sample paths of X . Assume ϕ is such that

$$(5.1) \quad \phi(X, t + A) = \phi(\theta_t X, A) \quad \text{for each } t, A.$$

Then N is stationary. This follows since

$$\theta_t N(\cdot) = \phi(X, t + \cdot) = \phi(\theta_t X, \cdot) \stackrel{d}{=} \phi(X, \cdot) = N(\cdot).$$

This result also holds if X is a stationary point process or a vector valued process. For instance, suppose X is a stationary pure-jump Markov process as in Example 4.13 or § 6.3. Consider the point process N of times that X jumps from some $x \in S$ to some $x' \in S'$ where $S \cap S' = \emptyset$. Then

$$N(A) = \phi(X, A) = \sum_{t \in A} 1(X_{t-} \in S, X_t \in S').$$

Clearly ϕ satisfies (5.1) and so N is stationary.

Another notion of stationarity is as follows. A point process N on \mathbb{R} has stationary intervals (or inter-point distances) if the sequence $\dots, W_{-1}, W_0, W_1, \dots$ is stationary: $(W_{n_1+h}, \dots, W_{n_k+h}) \stackrel{f}{=} (W_{n_1}, \dots, W_{n_k})$ for any n_1, \dots, n_k, h . If N is stationary, then intuition suggests that it has stationary intervals. This is true in only degenerate cases. Indeed, since the T_n 's are labeled such that $T_{-1} \leq T_0 < 0 \leq T_1$, then the distribution of W_1 will usually be different from $W_n, n \neq 1$: The W_1 is an interval covering a certain location 0 while the other W_n 's have no such restriction. This is called the waiting time paradox. For instance, if N is a Poisson process with rate λ , the W_1 is the sum of two independent exponential random variables $-T_0$ and T_1 both with mean λ^{-1} , while any other W_n is simply exponential with mean λ^{-1} . The Palm probabilities in § 5.3 shed light on this paradox.

5.2. Infinitesimal Properties

For the following discussion, we assume that N is a stationary point process on \mathbb{R} . It need not be simple. An application of Fubini's theorem (Chung (1972)) shows that $P\{N(\{t\}) = 0\} = 1$, for each t . Moreover, either $EN(I) < \infty$ for each finite interval I , or $EN(I) = \infty$ for each non-degenerate interval I . In either case,

$$EN_t = tEN_1, \quad t \geq 0,$$

where $EN_1 \leq \infty$. This follows since, by the stationarity of N ,

$$EN_{s+t} = EN_s + E(N_{s+t} - N_s) = EN_s + EN_t, \quad s, t \geq 0.$$

When N is a Poisson process with rate λ , then we know that

$$(5.2) \quad P\{N_t \geq 1\} = \lambda t + o(t) \quad \text{as } t \downarrow 0.$$

$$(5.3) \quad \lambda = EN_1.$$

$$(5.4) \quad P\{N_t \geq 2\} = o(t) \quad \text{as } t \downarrow 0.$$

Does each stationary process have an "intensity" λ as in (5.2)? If so, does λ always equal EN_1 ? A process satisfying property (5.4) is called orderly.

Lemma 1.7 ensures that if N is orderly, then it is simple and so one can order its point locations (they are "orderly"). When is N orderly? The following result addresses these issues.

Theorem 5.3. (a) (Khintchine) The limit $\lambda \equiv \lim_{t \rightarrow 0} P\{N_t \geq 1\}/t$ exists and

$$0 \leq \lambda \leq \infty.$$

(b) (Korolyuk) If N is simple, then $\lambda = EN_1 \leq \infty$.

(c) (Dobrushin) If N is simple and $EN_1 < \infty$, then N is orderly.

Proof (a) Clearly $f(t) \equiv P\{N_t \geq 1\}$ decreases to 0 as $t \downarrow 0$ and, for each $s, t \geq 0$,

$$f(s+t) = P\{N_s \geq 1\} + P\{N_s = 0, N_{s+t} \geq 1\} \leq f(s) + f(t).$$

Thus, λ exists by the following property of sub-additive functions: If $f: [0, b] \rightarrow \mathbb{R}$ is sub-additive ($f(s+t) \leq f(s) + f(t)$, $s, t \geq 0$) and $f(t) \rightarrow 0$ as $t \rightarrow 0$, then

$$\lim_{t \downarrow 0} f(t)/t = \sup_t f(t)/t \leq \infty.$$

(b) First note that since N is simple (its points are isolated), we can write

$$N_1 = \lim_{n \rightarrow \infty} S_n \quad \text{a.s., where } S_n \equiv \sum_{k=0}^n 1(N((k-1)/n, k/n] \geq 1) \text{ and } n \text{ runs through}$$

integers of the form 2^m . Then by Lebesgue's monotone convergence theorem and part (a), we have

$$EN_1 = \lim_{n \rightarrow \infty} ES_n = \lim_{n \rightarrow \infty} nf(1/n) = \lambda.$$

(c) Using the stationarity of N , we can write

$$\begin{aligned} EN_1 &= nEN(0, 1/n] \\ &= nf(1/n) + nP\{N(0, 1/n] \geq 2\} + n \sum_{k=3}^{\infty} P\{N(0, 1/n] \geq k\}. \end{aligned}$$

Letting $n \rightarrow \infty$, we know by (a) and (b) that $nf(1/n) \rightarrow EN_1$, and so the last two terms in the preceding display must tend to 0.

What can we say about the distribution of the number of points at a single location (or in a batch) when N is not simple? In case N is a compound Poisson process with rate λ and mass or batch-size distribution F , then

$$\begin{aligned} P\{1 \leq N_t \leq m\} &= \lambda F(m)t + o(t) \quad \text{as } t \downarrow 0, \text{ and} \\ F(m) &= \lim_{t \downarrow 0} P\{N_t \leq m \mid N_t \geq 1\}. \end{aligned}$$

These properties extend to the general stationary point process N as follows.

Theorem 5.4. For each m , the limit $\lambda_m \equiv \lim_{t \downarrow 0} P\{1 \leq N_t \leq m\}/t$ exists and $\lambda_m \uparrow$ some $\lambda \leq \infty$. When $\lambda < \infty$, then

$$(5.5) \quad F(m) \equiv \lambda_m / \lambda = \lim_{t \downarrow 0} P\{N_t \leq m \mid N_t \geq 1\}$$

and EN_1 is the mean of F .

Proof. The first assertion follows like Theorem 5.1 (a) since $g(t) \equiv P\{1 \leq N_t \leq m\}$ is sub-additive. The existence of the limit (5.5) follows by applying the first part and Theorem 5.1(a) to the conditional probability $P\{1 \leq N_t \leq m\}/P\{N_t \geq 1\}$. That EN_1 is the mean of F follows by an argument similar to that for Theorem 5.1 (b), which we omit.

Extensions 5.5. All of the preceding results hold under the weaker assumption that N is crudely stationary: $N(I+t) \stackrel{d}{=} N(I)$ for each t and interval I (see Chung (1972)). Korolyuk's result extends to non-stationary point processes that may also be on a general space. For this and other

insights into these infinitesimal properties, see Leadbetter (1972), Daley and Vere-Jones (1988) and their references.

5.3 Palm Probabilities

A Palm probability distribution of a point process is essentially a conditional probability distribution of the process given that it has a point at a certain location. Such probabilities shed light on inter-point distances. For instance, a stationary point process on \mathbb{R} will have stationary intervals with respect to its Palm probability even though it doesn't have stationary intervals with respect to the underlying probability measure.

We begin by defining Palm probabilities for non-stationary processes. Suppose N is a point process on $E = \mathbb{R}$ whose mean measure $\mu(A) = EN(A)$ is finite for bounded A . Let \mathcal{N} denote the set of all counting measures on E . The Campbell measure of N is a measure C on $E \times \mathcal{N}$ defined by

$$C(A \times M) = E[N(A)1(N \in M)], \quad \text{for } A, M \text{ Borel sets in } E, \mathcal{N}.$$

This measure admits the disintegration

$$(5.6) \quad C(A \times M) = \int_A P_t(M) \mu(dt),$$

where each P_t is a probability measure on \mathcal{N} (see §10 of Kallenberg(1983)).

The P_t , $t \in \mathbb{R}$, are called the Palm probability distributions of N . This definition is essentially that of Ryll-Nardzewski (1961) who gave a theoretical basis for earlier versions of this notion. Note that expression (5.6) is equivalent to

$$(5.7) \quad P_t(M) = C(dt \times M)/\mu(dt) = E[N(dt)1(N \in M)]/EN(dt).$$

We now restrict our attention to the case in which N is stationary and simple with $\lambda \equiv EN_1 < \infty$. Then from the right-hand side of (5.6), it follows

that $P_t(M) = P_0(\theta_{-t}M)$, where $\theta_{-t}M = \{\theta_{-t}\mu: \mu \in M\}$ and $\theta_{-t}\mu(A) \equiv \mu(A-t)$.

$A \in \mathcal{E}$. That is, the P_t 's are translations of P_0 . Accordingly, the single P_0 is called the Palm probability distribution of the stationary process N . In this case, the preceding definition of P_0 is expressed as

$$(5.8) \quad P_0(M) = (\lambda|A|)^{-1} E\left[\int_A 1((\theta_t N \subset M)N(dt)\right],$$

where $|A|$ is the area of $A \in \mathcal{E}$. The right-hand side is the same for each $A \in \mathcal{E}$, and so $A = (0,1]$ is a typical choice. Another expression for the last expectation is $\sum_n P\{\theta_{T_n} N \in M, T_n \in A\}$.

Following a common convention, we let $N^0 = \sum_n \delta_{T_n}$ denote another point process on \mathbb{R} , on some probability space, such that $P\{N^0 \in M\} = P_0(M)$ for each M . This probability measure upon which N^0 is based is not the same as the P we have been using for N . The distinction between these probabilities should be apparent from the events they measure. The N^0 is interpreted as " N conditioned that it has a point at 0" (the Palm version of N). This interpretation is justified by the property

$$P\{N^0(a,b] = m\} = \lim_{h \downarrow 0} P\{N(a,b] = m \mid N(-h, 0] \geq 1\}.$$

Furthermore, note that $P_0\{\mu: \mu\{0\} = 1\} = 1$ or $N^0\{0\} = 1$ a.s. This follows since using (5.8) with $A = (0,1]$ and $\theta_{T_n} N\{0\} = N\{T_n\} = 1$ a.s., we have

$$\begin{aligned} P\{N^0\{0\} = 1\} &= \lambda^{-1} \sum_n P\{\theta_{T_n} N\{0\} = 1, T_n \leq 1\} \\ &= \lambda^{-1} \sum_n P\{T_n \leq 1\} = \lambda^{-1} EN_1 = 1. \end{aligned}$$

Another feature of N^0 is that it has stationary intervals (even though N does not have stationary intervals). This statement is equivalent to $P_0(\theta_{T_1} M) = P_0(M)$, for each M , which one can prove using (5.8).

The following formulas relating P and P_0 are useful for proving results about N^0 or N .

Campbell Formula. For $\phi: \mathbb{R} \times \mathcal{N} \rightarrow \mathbb{R}_+$,

$$(5.9) \quad \lambda E[\phi(t, N^0)] = E\left[\int_{\mathbb{R}} \phi(t, \theta_t N) N(dt)\right],$$

provided these expectations exist. This follows from (5.5) when ϕ is a simple function and, for a general ϕ (which is a limit of simple functions), it follows by monotone convergence. A special case, for $f: \mathcal{N} \rightarrow \mathbb{R}_+$, is

$$\lambda E[f(N^0)] = E\left[\int_0^1 f(\theta_t N) dt\right].$$

Keep in mind that the expectation on the left is with respect to the probability for N^0 while the expectation on the right is with respect to the probability for N , which is different.

Expressing P in Terms of P_0 . For $f: \mathcal{N} \rightarrow \mathbb{R}_+$,

$$E[f(N)] = \lambda E\left[\int_0^{T_1^0} f(T_t N^0) dt\right] = \lambda E\left[\int_0^{-T_0^0} f(T_{-t} N^0) dt\right].$$

In particular,

$$P\{N \in M\} = \lambda E\left[\int_0^{T_1^0} 1(T_t N^0 \in M) dt\right].$$

A special case is the Palm-Khinchine formula

$$P\{N(0, t] = m\} = \lambda \int_t^\infty P\{N^0(0, s] = m\} ds.$$

Also, letting $F(t) = P\{T_1^0 \leq t\}$, we have (cf. Example 5.1)

$$(5.10) \quad \begin{aligned} P\{-T_0 > u, T_1 > v\} &= 1 - \lambda \int_0^{u+v} [1 - F(t)] dt \\ P\{-T_0 \leq t\} &= P\{T_1 \leq t\} = \lambda \int_0^t [1 - F(s)] ds. \end{aligned}$$

Our next result is a convenient formula for the Palm probability of a superposition of processes. Suppose N_1, \dots, N_n are independent, simple stationary point processes on \mathbb{R} with finite intensities $\lambda_1, \dots, \lambda_n$. Consider their superposition $N = N_1 + \dots + N_n$. Clearly N is stationary since it is of the form $N(A) = \phi(X, A)$ where $X \equiv (N_1, \dots, N_n)$ is stationary (recall Example 5.2), and N is simple with rate $\lambda \equiv \lambda_1 + \dots + \lambda_n$ since $N_j\{t\} = 0$ a.s. for each j, t and the N_1, \dots, N_n are independent.

Theorem 5.6. For $M \in \mathcal{N}$,

$$(5.11) \quad P\{N^0 \in M\} = \sum_{j=1}^n (\lambda_j / \lambda) P\{N_j^0 + \sum_{k \neq j} N_k \in M\},$$

where the N_j^0 , N_k , $k \neq j$ are independent.

Interpretation. This says that N^0 is distributed as $N_j^0 + \sum_{k \neq j} N_k$ with probability λ_j / λ , $j=1, \dots, n$ (i.e. N^0 is a "mixture" of these processes). The N_j^0 , N_k , $k \neq j$ on the right of (5.11) are defined on some probability space and their distributions are the same as the original processes with the same labels.

Proof. Two applications of the definition (5.8) and the stationarity of N_1, \dots, N_n yield

$$\begin{aligned} P\{N^0 \in M\} &= \lambda^{-1} E\left[\int_{(0,1]} 1(\theta_t N \in M) N(dt)\right] \\ &= \lambda^{-1} \sum_{j=1}^n E\left\{E\left[\int_{(0,1]} 1(\theta_t N_j + \sum_{k \neq j} \theta_t N_k \in M) N_j(dt) \mid N_k, k \neq j\right]\right\} \\ &= \lambda^{-1} \sum_{j=1}^n \lambda_j E[1(N_j^0 + \sum_{k \neq j} \theta_t N_k \in M)] \\ &= \lambda^{-1} \sum_{j=1}^n \lambda_j P\{N_j^0 + \sum_{k \neq j} N_k \in M\} / \end{aligned}$$

Example 5.7. Suppose N_1, \dots, N_n are independent renewal processes with waiting time distributions F_1, \dots, F_n that have means $\lambda_1^{-1}, \dots, \lambda_n^{-1}$. Their superposition $N = N_1 + \dots + N_n$ is stationary, but it will generally not be a stationary renewal process. Theorem 5.6 and (5.10), however, tell us that the T_0 and T_1 for N are such that

$$P\{-T_0 > u, T_1 > v\} = \sum_{j=1}^n (\lambda_j / \lambda) \bar{F}_j(u) \bar{F}_j(v) \prod_{k \neq j} \bar{G}_k(u+v)$$

where $G_k(t) = \lambda_k \int_0^t [1 - F_k(s)] ds$ and $\bar{F}(t) = 1 - F(t)$.

6. POINT PROCESSES CHARACTERIZED BY MARTINGALES

We now discuss point processes on R_+ whose evolution is characterized by an increasing history of observed events or information. Here are some motivating examples. Suppose $N = \{N_t : t \in R_+\}$ is a Poisson process on R_+ with mean measure $A_t = \int_0^t \lambda_s ds$ (following the convention of this area, we now use A instead of Λ). One can interpret N as evolving over time and its evolution or dynamics are given by

$$(6.1) \quad E[dN_t | F_{t-}] = P\{dN_t = 1 | F_{t-}\} = \lambda_t dt + o(dt),$$

where $dN_t = N_t - N_{t-}$, $F_t \equiv \sigma(N_s, s \leq t)$, the σ -field of the history of N up to time t , and F_{t-} is the history on $[0, t)$. A more concise way of expressing (6.1) is to say that the process $M_t \equiv N_t - A_t$, $t \in R_+$, is an F_t -martingale. That is, $E[M_t | F_s] = M_s$, $s \leq t$, which is equivalent to

$$(6.2) \quad E[N_t - N_s - \int_s^t \lambda_u du | F_s] = 0.$$

As another example, suppose that N is a Cox process directed by $A_t = \int_0^t \lambda_s ds$ where λ is a non-negative stochastic process. Then the evolution of N is also characterized by (6.1), (6.2), where λ_t is random and $F_t \equiv \sigma(N_s, s \leq t; \lambda_u, u \geq 0)$, the σ -field of N up to t and the entire trajectory of λ as well. For the final example, suppose that $\{X_t; t \in R_+\}$ is a birth and death queueing process with state space $\{0, 1, \dots\}$ and state-dependent arrival and service rates $\alpha(n)$, $\beta(n)$. That is, when $X_t = n$, the time to the next potential arrival is exponential with mean $\alpha(n)^{-1}$ and the time to the next potential service completion is exponential with mean $\beta(n)^{-1}$. Consider the point process N_t of the number of customer arrivals up to time t . Again, the evolution of N is characterized by (6.1), (6.2) with $\lambda_t \equiv \alpha(X_{t-})$ and $F_t \equiv \sigma(X_s; s \leq t)$, the history of X up to t , which also includes that of N up

to t . This is sometimes referred to as a state-dependent Poisson process with random intensity λ_t (be careful with this loose terminology since N is generally not a Poisson or even a Cox process).

Such examples motivated the development of a general theory of point processes with dynamics as above. The link between these point processes and their associated martingales has led to a martingale calculus of point processes, which is part of the modern stochastic calculus that deals with integration with respect to Wiener processes, martingales or semi-martingales. Point processes on R_+ are special submartingales or semi-martingales.

Much of the current mathematical research on point processes deals with this class of processes. Little is known about these processes on R^2 or other partially ordered spaces; see for instance Merzbach and Nualart (1986). There are a number of results in this area, such as in filtering and optimal dynamic control, that have potential applications in operations research. Basic references are Bremaud (1981), Liptser and Shiriyayev (1978), Ikeda and Watanabe (1981), and Karr (1986). Unfortunately, we cannot get into these lengthy topics. We will be content with introducing the basic notion of a compensator and showing how it is used in Poisson limit theorems and approximations.

6.1 Compensators of Point Processes

As one would expect, the stochastic process A_t appearing in the martingale $M_t = N_t - A_t$ in the examples above plays an important role in characterizing N . Such processes, called F_t -compensators of N , are the subject of this subsection.

Let (Ω, F, P) be a probability space. This will be the underlying space for all of our processes. Let $\{F_t: t \in R_+\}$ be a filtration or history on (Ω, F) : a family of sub- σ -fields of F that are increasing ($F_s \subset F_t, s \leq t$). The F_t represents the information one observes up to time t . We assume, as usual, that F_0 contains all P -null events and that F_t is right-continuous ($F_t = \bigcap_{u>t} F_u$). Suppose $X = \{X_t: t \in R_+\}$ is a real-valued stochastic process on (Ω, F, P) . The internal history of X is $F_t^X = \sigma(X_s: s \leq t)$, the σ -field of the events of X up to time t . The X is F_t -adapted, if $F_t^X \subset F_t$ for each t . The X is F_t -predictable if it is F_t -adapted and each set $\{(t, \omega) \in R_+ \times \Omega: X_t(\omega) \leq x\}$ is in the smallest σ -field on $R_+ \times \Omega$ that contains the sets $(s, t] \times B, s \leq t, B \in F_s$. For our purposes, one can define X as being F_t -predictable if it is F_t -adapted and has left-continuous paths (such processes form a large class of predictable processes). The process X is increasing if $X(0) = 0$ and its sample paths are nondecreasing and right-continuous.

Consider a point process $N = \{N_t: t \in R_+\}$ defined on (Ω, F, P) that is F_t -adapted and has point locations $T_1 < T_2 < \dots$ with $T_n \rightarrow \infty$ a.s. For convenience, we assume that each N_t has a finite mean (this allows us to use martingales instead of more general local martingales). Then there is an increasing F_t -predictable process $A = \{A_t, t \in R_+\}$ such that, for each F_t -predictable process $C = \{C_t: t \in R_+\}$,

$$(6.3) \quad E \left[\int_0^\infty C_t dN_t \right] = E \left[\int_0^\infty C_t dA_t \right].$$

The A is unique up to P -null events. This process A is the F_t -compensator (or dual predictable projection) of N . We sometimes call N an F_t -point process with compensator A . The condition (6.3) holds if and only if the process $M_t = N_t - A_t, t \in R_+$, is a martingale. The M is the process N

compensated by A . The representation $N_t = M_t + A_t$ is the Doob-Meyer decomposition of N (viewed as a submartingale). Keep in mind that A depends on the choice of F_t ; the smallest possible F_t is N 's internal history F_t^N .

The A often has the form $A_t = \int_0^t \lambda_s ds$, where $\lambda = \{\lambda_t; t \in R_+\}$ is a non-negative F_t -predictable process called the F_t -stochastic intensity of N . This intensity has the interpretation (6.1) and, in many cases,

$$\lambda_t = \lim_{h \rightarrow 0} E[N_{t+h} - N_t | F_t] / h.$$

The process N is an F_t -Poisson process if, for each $s \leq t$, $N_t - N_s$ is a Poisson random variable independent of F_s . The N has F_t -independent increments if for each $s \leq t$, $N_t - N_s$ is independent of F_s .

The following are some basic properties of compensators:

- (i) N is an F_t -Poisson process if and only if A is deterministic and continuous. In this case $E[N_t - N_s | F_s] = A_t - A_s$, $s \leq t$.
- (ii) N has F_t -independent increments if and only if A is deterministic.
- (iii) N has F_t -conditionally independent increments given \mathcal{G} , a sub- σ -field of F_0 , if and only if A is a \mathcal{G} -measurable function.
- (iv) If A is a.s. continuous, then N is stochastically continuous ($P\{dN_t = 1\} = 0$, $t \in R_+$).
- (v) If X is an F_t -predictable process with $\int_{[0,t]} |X_s| dA_t < \infty$ a.s., $t \in R_+$, then the process $Z_t = \int_{[0,t]} X_s dM_s$, $t \in R_+$ is an F_t -martingale.

We have already seen three examples of compensators of point processes in the introduction. Here are two more.

Example 6.1. Renewal Processes and Some Relatives. Suppose N is as above with $F_t = \mathcal{G} \vee F_t^N$ where \mathcal{G} is some σ -field (e.g., for the Cox process above, $\mathcal{G} = \sigma(\lambda_s; s \in R_+)$). Let $F_n(x) = P\{T_{n+1} - T_n \leq x | F_n\}$. Then the F_t -compensator of N is

$$A_t = A_{T_n} + \int_0^{t-T_n} [1 - F_n(x-)]^{-1} F_n(dx), \quad T_n \leq t \leq T_{n+1}.$$

This resembles a hazard rate in reliability since $dA_{T_n} = dF_n(t)/[1 - F_n(t-)]$.

In particular, if N is a renewal process with waiting time distribution F and $F_t = F_t^N$, then the compensator is as above with the random F_n replaced by F .

Example 6.2. Jump Times of Markov Processes. Suppose $X = \{X_t; t \in \mathbb{R}_+\}$ is a Markov process with countable state space X and transition rates $q(x,y)$ (as in Example 4.13). We assume that $0 < \sum_{y \neq x} q(x,y) < \infty$ and that X cannot take

an infinite number of jumps in a finite time interval. A variety of point processes associated with jumps of X can be modeled as follows. Consider the point process N of times at which X jumps from some state x to another state y where (x,y) are in the set $J \subset X \times X$ and J does not contain pairs (x,x) .

That is,

$$N_t = \sum_{s \leq t} 1((X_{s-}, X_s) \in J), \quad t \in \mathbb{R}_+.$$

Then an easy check shows that the F_t^X -compensator of N is

$$A_t = \int_0^t \sum_{(x,y) \in J} q(x,y) 1(X_{s-} = x) ds, \quad t \in \mathbb{R}_+.$$

The X_{s-} may be replaced by X_s since X is stochastically continuous. These point processes are useful for modeling flows of customers in queueing networks as we discuss later.

Compensators for marked point processes are defined similarly. Suppose $N = \sum_{n=1}^{\infty} \delta_{T_n, Z_n}$ is a marked point process on $\mathbb{R}_+ \times E$ such that the point process $N_t = \sum_n 1(T_n \leq t)$ is as above. Then the compensator of N is the unique random measure A on $\mathbb{R}_+ \times E$ such that $A_t(B) \equiv A([0, t] \times B)$, $t \in \mathbb{R}_+$, is F_t -predictable for each $B \subset E$ and, for each predictable process $C = \{C(t, x); (t, x) \in \mathbb{R}_+ \times E\}$,

$$E \left[\int_{R_+ \times E} C(t, x) N(dt dx) \right] = E \left[\int_{R_+ \times E} C(t, x) A(dt dx) \right].$$

6.2 Poisson Convergence and Approximations.

Many properties of a point process on R_+ can be expressed in terms of its compensator. For instance, one might expect that a sequence of point processes would converge in distribution if their compensators converge appropriately. One such Poisson limit theorem is as follows. This and other limit theorems for processes with independent increments or conditionally independent increments appear in Kabanov et. al. (1983), and Kabanov and Liptser (1983) (they also discuss marked point processes).

Theorem 6.3. Let N^n be an F_t^n -point process on R_+ with compensator A^n , $n=1,2,\dots$, and let N be an F_t -Poisson process with (deterministic continuous) compensator A . If $A_t^n \xrightarrow{d} A_t$ for each t , then $N^n \xrightarrow{d} N$.

The following total variation bounds are useful for analyzing rates of convergence in the preceding setting or for establishing Poisson approximations. Suppose N' is a point process with F_t' -intensity λ_t' and N is a Poisson process with F_t -intensity λ_t . Let P_t' and P_t denote the probability distributions of the respective processes N' and N on the interval $[0, t]$ (e.g. $P_t(B) = P\{N \in B\}$ where B is a Borel subset of sample paths of N on $[0, t]$). The total variation distance between P_t' and P_t (recall §3.4) is defined by $d(P_t', P_t) = \sup_B |P_t'(B) - P_t(B)|$.

Total Variation Bounds 6.4.

$$(i) \quad d(P_t', P_t) \leq E \int_0^t |\lambda_s' - \lambda_s| ds.$$

(ii) If $\lambda_s = E\lambda_s'$, then

$$d(P_t', P_t) \leq \int_0^t \text{Var } \lambda_s'^{1/2} ds.$$

(iii) Using the notation of §3.4,

$$d(N'_t, N_t) = 1/2 \sum_{n=0}^{\infty} |P\{N'_t = n\} - P\{N_t = n\}| \leq d(P'_t, P_t).$$

Inequality (i) is proved in Brown (1983) and Kabanov et. al (1983).

Inequality (ii) follows from (i) since $E |\lambda'_S - E\lambda'_S| \leq \text{Var } \lambda'_S^{1/2}$. And (iii) follows from the definitions.

Example 6.5. Poisson Approximations of Jump Times of Markov Processes.

Suppose, as in Example 6.2, that X is a Markov process and N' is the point process of jumps of X from x to y where $(x,y) \in J$. We observed that N' has

F_t^X -intensity $\lambda'_t = \sum_{(x,y) \in J} q(x,y) 1\{X(t) = x\}$. Assume that X is stationary

with distribution $\pi(x) = P\{X(t) = x\}$. Let N be a Poisson process with

intensity $\lambda_t = E\lambda'_t = \sum_{(x,y) \in J} q(x,y)\pi(x)$. Then from Bound 6.4 (i), we have

$$\begin{aligned} d(P'_t, P_t) &\leq t \sum_{(x,y) \in J} q(x,y) E|1\{X(t) = x\} - \pi(x)| \\ &= 2t \sum_{(x,y) \in J} q(x,y)\pi(x)[1 - \pi(x)]. \end{aligned}$$

When the last sum is small, then N' is close to being Poisson. A specific illustration follows.

6.3 Customer Flows in a Jackson Queueing Network That are Approximately Poisson.

We now assume that the Markov process above is an open Jackson queueing network process defined as follows. Consider a network of J nodes representing service stations. Customers enter the nodes $1, \dots, J$ from outside the network according to independent Poisson processes with respective rates ν_1, \dots, ν_J . Each node j operates as an isolated single server whose service times are independent exponential random variables with mean ϕ_j^{-1} . Customers are served one at a time, under any priority scheme. A customer, after being served at node j , goes immediately to node k with

probability p_{jk} , $k=1, \dots, j$, or exits the network with probability p_{j0} ($\sum_{k=0}^J p_{jk} = 1$). Let $X(t) = (X_1(t), \dots, X_J(t))$ denote the numbers of customers at the respective nodes at time t . Then X is a Markov process with state space $\mathcal{X} = \{x = (x_1, \dots, x_J) : x_j = 0, 1, \dots\}$ and transition rates

$$q(x, y) = \begin{cases} v_j & \text{for } y = x + e_j \\ \phi_j p_{jk} & y = x - e_j + e_k, \quad x_j \geq 1 \\ \phi_j p_{j0} & y = x - e_j \end{cases}$$

and $q(x, y) = 0$ elsewhere, where e_j is the vector with 1 in position j and 0's elsewhere.

Consider the routing probabilities $\{p_{jk}\}$ $j, k=0, \dots, J$ as a Markov matrix, where $p_{0k} = v_k / \sum v_p$. Assume that this matrix is irreducible. Then there is a unique solution $\alpha_1, \dots, \alpha_J$ to the so-called traffic equations

$$\alpha_j = v_j + \sum_{k=1}^J v_k p_{kj}, \quad j=1, \dots, J.$$

Namely, $\alpha_0 = q_0 / \sum v_p$, $\alpha_j = q_j$ ($j \geq 1$) where q_0, \dots, q_J is the probability distribution satisfying $q_j = \sum_{k=0}^J q_k p_{kj}$, $j=0, \dots, J$. This assumption ensures that there is a positive probability that each customer may reach any node. We also assume that the traffic intensity $\rho_j \equiv \alpha_j / \phi_j < 1$ for each j . Thus

the network process X has the stationary distribution $\pi(x) = \prod_{j=1}^J (1 - \rho_j) \rho_j^{x_j}$. Finally, assume that X is stationary.

Now, consider the point process

$$N_{i,j}(t) = \sum_{s \leq t} \sum_x 1(X(s-) = x, X(s) = x - e_i + e_j), \quad (i, j \in \mathcal{C}_+)$$

of the number of times customers move from node i to node j in the time

interval $[0, t]$. Similarly, let $N_{i0}(t)$ denote the number of times customers exit the network from node i up to time t . Also consider the process

$$N_j(t) = \sum_{s \leq t} \sum_x 1(X(s-) = x, X(s) = x + e_j \text{ or } x - e_i + e_j \text{ for some } i), \quad t \in \mathbb{R}_+,$$

of the number of customer arrivals into node j up to time t . From Example 6.2, we know that these processes have the respective F_t^X -intensities

$$\lambda_{ij}(t) = 1(X_i(t) \geq 1) \phi_i p_{ij}, \quad \text{and } \lambda_j(t) = \nu_j + \sum_i 1(X_i(t) \geq 1) \phi_i p_{ij}.$$

Since X is stationary, it follows that N_{ij} and N_j are stationary with rates

$$E \lambda_{ij}(t) = \rho_i \phi_i p_{ij} = \alpha_i p_{ij}$$

$$E \lambda_j(t) = \nu_j + \sum_i \rho_i \phi_i p_{ij} = \alpha_j.$$

It is well known that N_{10}, \dots, N_{J0} are independent Poisson processes (in the usual sense when only their internal histories are observed and X is not) with the respective rates $\alpha_1 p_{10}, \dots, \alpha_J p_{J0}$. This follows using filtering or a reversibility argument; see for instance Bremaud (1981). Similarly, one can show that if I, J are two disjoint subsets of nodes such that from any node in J a customer cannot reach a node in I , then $\{N_{ij}\}$, $i \in I$, $j \in J$, are independent Poisson processes with respective rates $\{\alpha_i p_{ij}\}$.

We now consider N_{ij} as an F_t^X -point processes (i.e. X as well as N_{ij} is observed). Our interest is in how close N_{ij} is to being Poisson. Let d_{ij}^t denote the total-variation distance between the distribution of N_{ij} and that of a Poisson process with rate $\alpha_i p_{ij}$ on $[0, t]$. Then by the Bound 6.4 (i), we have

$$d_{ij}^t \leq \int_0^t E |\lambda_{ij}(s) - E \lambda_{ij}(s)| ds$$

$$= t \phi_i p_{ij} E |1(X_i(t) \geq 1) - \rho_i|$$

$$= 2t \phi_i p_{ij} \rho_i (1 - \rho_i) = 2t \alpha_i p_{ij} (1 - \rho_i).$$

Thus N_{ij} will be approximately Poisson when the traffic into i is light ($\rho_i \cong 0$), the traffic into i is heavy ($\rho_i \cong 1$), or the traffic between i and j is light ($\alpha_i p_{ij} \cong 0$).

Similarly, let d_j^t denote the total-variation distance between the distribution of N_j and that of a Poisson process with rate α_j on $[0, t]$. Then, as above, $d_j^t \leq 2t \sum_i \phi_i p_{ij} \rho_i (1 - \rho_i)$. Consequently, N_j will be approximately Poisson under the conditions above or when the network is large and the dispersion of customers via the p_{ij} 's is relatively even (the p_{ij} 's are approximately equal and small). Brown and Pollett (1982) discuss this approximation and related ones for closed networks (they sometime use the looser Bound 6.4(ii) instead of (i) for convenience).

References

- Baccelli, F. and P. Brémaud (1987). Palm Probabilities and Stationary Queues. Lect. Notes Statist. 41, Springer-Verlag, New York.
- Bartfai, P. and Tomkó, J. (1981). Point Processes and Queueing Problems. North-Holland, Amsterdam.
- Bateman, H. (1910). Note on the probability distribution of α particles. Phil. Mag. (6) 20, 704-707.
- Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.
- Brémaud, P. (1981). Point Processes and Queues: Martingale Dynamics. Springer-Verlag, New York.
- Brown, T.C. (1983). Some Poisson approximations using compensators. Ann. Probability 11, 726-744.
- Brown, T.C. and P.K. Pollett (1982). Some distributional approximations in Markovian queueing networks. Adv. Appl. Prob. 14, 654-671.
- Chung, K. (1972). Crudely stationary point processes. Amer. Math. Monthly 79, 867-877.
- Cinlar, E. (1972). Markov additive processes, I,II. Z. Wahrscheinlichkeitstheorie Verw. Geb. 24, 85-93, 85-101.
- Cinlar, E. (1975). Introduction to Stochastic Processes. Prentice-Hall, Englewood Cliffs, NJ.
- Cox, D.R. (1955). Some statistical methods connected with series of events. J.R. Statist. Soc. B 17, 129-164.
- Cox, D.R. and P.A.W. Lewis (1966). The Statistical Analysis of Series of Events. Methuen, London.
- Cox, D. and V. Isham (1980). Point Processes. Chapman and Hall, London.
- Cramér, H. and M.R. Leadbetter (1967). Stationary and Related Stochastic Processes. Wiley, New York.
- Daley, D.J. and Vere-Jones, D. (1972). A summary of the theory of point processes. In Stochastic Point Processes, P.A.W. Lewis Ed. Wiley, New York.
- Daley, D.J. and Vere-Jones, D. (1988). Forthcoming book on point processes and random measures.
- Dale, P.J. and R.K. Milne (1983). Negative binomial quadrat counts and point processes. Scand. J. Statist. 10, 257-267.

- Elliott, R.J. (1982). Stochastic Calculus and Applications. Springer-Verlag, New York.
- Erlang, A.K. (1909). The theory of probabilities and telephone conversations. Nyt. Tidsskr. Mat. B 20, 33-41.
- Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol 2, 2nd ed. Wiley, New York.
- Franken, P., König, D., Arndt, V. and V. Schmidt (1981). Queues and Point Processes. Akademie-Verlag, Berlin.
- Gibbs, J.W. (1902). Elementary Principles in Statistical Mechanics. Yale University Press.
- Grandell, J. (1976). Doubly Stochastic Poisson Processes. Lect. Notes Math., 529, Springer-Verlag, New York.
- Grandell, J. (1977). Point processes and random measures. Adv. Appl. Prob. 9, 502-526.
- Graunt, J. (1662). Observations on the London Bills of Mortality.
- Gregoire, G. (1983). Negative binomial distributions for point processes. Stoch. Proc. Appl. 16, 179-188.
- Grigelionis, B. (1963). On the convergence of sums of random step processes to a Poisson process. Theory Prob. Appl., 8, 177-1182.
- Gut, A. and N.U. Prabhu (1987). Random Walks, Renewal Theory and Regenerative Phenomena. (Preprint).
- Hinderer, K. (1985). A unifying method for some computations in renewal theory. Zeitschr. Angew. Math. Mech.
- Hinderer, K. (1986). Remarks on direct Riemann integrable functions.
- Hsing, T., Hüsler, J. and M.R. Leadbetter (1986). On the exceedances point process for a stationary sequence. (Preprint).
- Iked, N. and S. Watanabe (1981). Stochastic Differential Equations and Diffusion Processes. North-Holland, Amsterdam.
- Jacobsen, M. (1982). Statistical Analysis of Counting Processes. Lect. Notes Statist. 12, Springer-Verlag, New York.
- Jagers, P. (1972). Aspects of random measures and point processes. in Stochastic Point Processes, P.A.W. Lewis ed. Wiley, New York
- Kabanov, Yu. M. and R. Sh. Liptser (1983). On convergence in variation of the distributions of multivariate point processes. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.

- Kabanov, Yu. M., Liptser, R. Sh. and A.N. Shirayev (1983). Weak and strong convergence of the distribution of counting processes. Theor. Probability Appl. **28**, 303-336.
- Kallenberg, O. (1983). Random Measures, 3rd ed. Akademie-Verlag, Berlin/Academic Press, New York.
- Karr, A.F. (1986). Point Processes and Their Statistical Inference. Dekker, New York.
- Kendall, D.G. (1974). Foundations of a theory of random sets. In Stochastic Geometry. Wiley, New York.
- Kesten, H. (1974). Renewal theory for functionals of a markov chain with general state space. Ann. Probability 355-386.
- Kesten, H. (1982). Percolation Theory for Mathematicians. Birkhauser, Boston.
- Kienderman, R. and J.L. Snell (1980). Markov Random Fields and Their Applications. American Math. Soc. Providence.
- Khinchin, A. Ya. (1960). Mathematical Methods in the Theory of Queueing. Griffin, London.
- Kingman, J.F.C. (1972). Regenerative Phenomena. Wiley, New York.
- Leadbetter, M.R. (1972). On basic results of point process theory. Proc. Sixth Berkeley Symp. Math. Statist. Prob. **3**, 449-462.
- Leadbetter, M.R., Lingren, G. and H. Rootzén (1983). Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, New York.
- Lewis, P.A.W. (ed.) (1972). Stochastic Point Processes. Wiley, New York.
- Lewis, P.A.W. and G.S. Shedler (1979). Simulation of nonhomogeneous Poisson processes by thinning. Naval Res. Logist. Quart. **29**, 403-413.
- Liggett, T.M. (1985). Interacting Particle Systems. Springer-Verlag, New York.
- Liptser, R.Sh. and A.N. Shirayev (1978). Statistics of Random Process. Vols. I and II. Springer-Verlag, New York.
- Lotka, A.J. (1939). A contribution to the theory of self-renewing aggregates, with special reference to industrial replacement. Ann. Math. Statist. **10**, 1-25.
- Matheron, G. (1975). Random Sets and Integral Geometry. Wiley, New York.
- Matthes, K., Kerstan, J. and J. Mecke (1978). Infinitely Divisible Point Processes. Wiley, Chichester (German ed. 1974).

- Merzbach, E. and D. Nualart (1986). A characterization of the spatial Poisson process and changing time. Ann. Probability, **14**, 1380-1390.
- Neuts, M.F. (1979). A versatile Markovian point process. J. Appl. Prob. **16**, 764-779.
- Neveu, J. (1977). Processus Ponctuels. Lect. Notes Math. **598**, 249-447, Springer-Verlag, New York.
- Neyman, J. and Scott, E.L. (1958). Statistical approach to problems of cosmology (with discussion). J. Roy. Statist. Soc. B **20**, 1-43.
- Palm, C. (1943). Intensitätsschwankungen in Fernspreverkehr. Ericsson Techniques **44**, 1-189.
- Poisson, S.D. (1837). Recherches sur la Probabilité des Judgements en Matière Criminelle et en Matière Civile. Précédies des Rigles Cénréales du Calcul des Probabilités. Bachelier, Paris.
- Rényi, A. (1967). Remarks on the Poisson process. Studies Sci. Math. Hungar. **2**, 119-123.
- Ripley, B.D. (1976). Locally finite random sets: foundations for point process theory. Ann. Probability **4**, 983-994.
- Ripley, B.D. (1981). Spatial Statistics. Wiley, New York.
- Rolski, T. (1981). Stationary Random Processes Associated with Point Processes. Lect. Notes Statist. **5**. Springer-Verlag, New York.
- Rozanov, Yu. A. (1980). Markov Random Fields. Springer-Verlag, New York.
- Ryll-Nardzewski, C. (1961). Remarks on processes of calls. Proc. Fourth Berkeley Symp. Math. Statist. Prob. **2**, 455-465.
- Serfling, R.J. (1975). A general Poisson approximation theorem. Ann. Probability **3**, 726-731.
- Serfozo, R.F. (1980). High level exceedances of regenerative and semi-stationary processes. J. Appl. Prob. **17**, 423-431.
- Serfozo, R.F. (1982). Convergence of Lebesgue integrals with varying measures. Sankhyā, **44**, Ser. A 380-402.
- Serfozo, R.F. (1984a). Thinning of cluster processes: convergence of sums of thinned point processes. Math. of Operations Res. **9**, 522-533.
- Serfozo, R.F. (1984b). Rarefactions of compound point processes. J. Appl. Prob. **21**, 710-719.
- Serfozo, R.F. (1986). Compound Poisson approximations for sums of random variables. Ann. Probability **14**, 1391-1398.

Snyder, D.L. (1975). Random Point Processes. Wiley-Interscience, New York.

Weibull, W. (1939). A statistical theory of the strength of materials. Ing. Velensk. Akad. Hanal. Stockholm #151.

Wold, H. (1948). On stationary processes and Markov chains. Skand. Aktuar. 31, 229-240.

END

11-87

DTIC